

Ett likvärdigt betygssystem

Volym 2

*Betänkande av Utredningen om
likvärdiga betyg och meritvärden*

Stockholm 2025



STATENS OFFENTLIGA
UTREDNINGAR

SOU 2025:18

SOU och Ds finns på regeringen.se under Rättsliga dokument.

Svara på remiss – hur och varför
Statsrådsberedningen, SB PM 2021:1.

Information för dem som ska svara på remiss finns tillgänglig på regeringen.se/remisser.

Layout: Kommittéservice, Regeringskansliet

Omslag: Elanders Sverige AB

Tryck och remisshantering: Elanders Sverige AB, Stockholm 2025

ISBN 978-91-525-1155-8 (tryck)

ISBN 978-91-525-1156-5 (pdf)

ISSN 0375-250X

Innehåll

Bilaga 1	Kommittédirektiv 2023:95	9
Bilaga 2	Betyg, bedömning och centrala prov i ett urval av länder samt International Baccalaureate Diploma Program (IB).....	19
	Referenser	97
Bilaga 3	Statistiska analyser	113
1.1	Meritvärdeskalibrering – några exempel enligt en linjär modell	113
1.1.1	Den linjära kalibreringsmodellen.....	113
1.1.2	Hypotetiska förenklade exempel.....	114
1.2	Konsekvenser av meritvärdeskalibrering grundskolan	123
1.2.1	Data och metod	124
1.2.2	Konsekvenser för hela elevpopulationen och betygssystemet som helhet	129
1.2.3	Konsekvenser för olika undergrupper av elever ..	134
1.2.4	Nettoeffekter på olika elevgrupper – regressionsanalys	145
1.2.5	Förändringar i elevsammansättning för den högst presterande decilen (topp 10 %) av elevkohorten i grundskolan	148
1.3	Konsekvenser av meritvärdeskalibrering i gymnasieskolan	151
1.3.1	Data och metod	151
1.3.2	Kalibreringssteget.....	155
1.3.3	Meritvärdesberäkning.....	156
1.3.4	Konsekvenser för hela elevpopulationen och betygssystemet som helhet	156

1.3.5	Konsekvenser för olika undergrupper av elever..	164
1.3.6	Regressionsanalys – systematiska effekter	174
1.3.7	Elevsammansättning för den högsta decilen av eleverna utifrån meritvärde och kalibrerat meritvärde.....	176
1.3.8	Konsekvenser för antagning till en specifik och attraktiv högskoleutbildning	181
1.3.9	Konsekvenser av ”Strannegårdskriteriet”	188
1.4	Meritvärdeskalibrering vid specialfall.....	190
1.4.1	Små skolor	191
1.4.2	Elever som inte deltar på samtliga prov	194
1.5	Simulering av betygsmodeller – uppdaterade resultat.....	198
1.5.1	Grundläggande principer för simulering	199
1.5.2	Betygsmodeller.....	200
1.5.3	Resultat	201
1.5.4	Hur kan dessa resultat översättas till en meritvärdeskalibreringsmodell?	207
1.6	Skalning av provresultat, betyg och meritvärden	208
1.6.1	Möjlighet att sprida ut elevers provgenomförande på flera år.....	208
1.7	Vilken skala/skalor är lämpligast för återrapportering av betyg, betygsvärde, provresultat respektive meritvärde? ...	213
1.7.1	Provskala.....	213
1.7.2	Meritvärdesskala.....	214
1.8	Korrelationer (samband) mellan resultat i ämnen med nationella prov och ämnen utan nationella prov.....	215
1.8.1	Data och metod.....	215
1.8.2	Resultat och slutsatser	216
1.9	Alternativa kalibreringsmodeller – Ekvipercentilkalibrering.....	217
1.9.1	De grundläggande skillnaderna mellan kvartil- och linjär modell.....	218
1.9.2	Kvartilmodellen steg för steg	219
1.9.3	Kvartilmodellen vs den linjära modellen – ett exempel	223

Bilaga 4	Elever och lärares perspektiv på mer likvärdiga betyg och meritvärden.....	229
Bilaga 5	Improving the Comparability and Reliability of the Sweden National Tests	243
Bilaga 6	Underlagsrapport till utredningen: Likvärdiga betyg och meritvärden (Dir. 2023:95)	265
Bilaga 7	Konsekvenser av låga och underkända betyg för elevers motivation, prestationer och välbefinnande ...	285
Bilaga 8	Betydelsen av en skarp godkäntgräns i betygssystemet.....	309
Bilaga 9	Prövningar i grund- och gymnasieskolan.....	359

Statens offentliga utredningar 2025

Kommittédirektiv 2023:95

Likvärdiga betyg och meritvärden

Beslut vid regeringssammanträde den 22 juni 2023

Sammanfattning

En särskild utredare ska analysera och lämna förslag på förändringar i betygssystemet och systemet för meritvärdering. Förslagen ska syfta till att betyg och meritvärden från grundskolan och gymnasial nivå mer rättvisande speglar elevers ämneskunskaper samt till att motverka betygsinflation.

Utredaren ska bl.a.

- analysera för- och nackdelar med systemet för betygssättning och meritvärdering i det svenska skolväsendet i relation till andra länders system, till av Skolverket presenterade modeller och till av utredaren vidareutvecklade eller framtagna modeller,
- lämna ett huvudförslag, samt ett alternativt förslag på en systemförändring som innebär att resultat på nationella prov, examensprov, eller liknande, kan användas för att öka likvärdigheten i betyg eller meritvärden,
- lämna förslag på hur betygsinflation och bristande likvärdighet i betygssättning kan motverkas även i ämnen som inte har nationella prov, och
- lämna nödvändiga författningsförslag.

Uppdraget ska redovisas senast den 21 februari 2025.

Uppdraget att föreslå ett system för mer likvärdiga betyg och meritvärden

De betyg som sätts i skolan ska spegla elevers kunskaper på ett rättvisande sätt

Betygen inom det svenska skolsystemet sätts i huvudsak av den undervisande läraren (3 kap. 16 § skollagen [2010:800]). Till stöd för betygssättningen har läraren en betygsskala med nationellt formulerade och framtagna betygskriterier för varje betygssteg (se t.ex. 10 kap. 17, 19 och 20 §§ skollagen samt 6 kap. 7 § skolförordningen [2011:185]). Betygskriterierna ska bidra till likvärdighet i lärarnas bedömning av elevernas kunskaper. Vid betygssättning i en kurs eller ett ämne där det ges ett nationellt prov ska elevens resultat på det provet särskilt beaktas om det inte finns särskilda skäl (se t.ex. 10 kap. 20 a § skollagen). Detta innebär att de nationella proven är betygsstödjande på individnivå. Det är dock fortfarande den betygssättande läraren som avgör exakt hur provresultaten beaktas vid betygssättningen.

I Sverige används betyg både när man söker arbete och för tillträde och urval till utbildning. För behörighet till gymnasieskolans nationella program ställs krav på godkända betyg i ett visst antal ämnen och i några särskilt utpekade ämnen. Motsvarande gäller för behörighet till högskoleutbildning. Sådana krav på förkunskaper har som syfte att säkerställa att den som antas till utbildning på en högre nivå har de kompetenser som behövs för att kunna tillgodogöra sig utbildningen. Till många utbildningar konkurrerar ungdomar dessutom med sina betyg. Det är därför av stor vikt för individen att betygssättning sker på så likvärdiga grunder som möjligt, dvs. att ett visst betyg avspeglar samma kunskapsnivå, oavsett vilken skola eleven gått på eller vilken lärare eleven har haft. Om betygssättningen inte är likvärdig mellan lärare eller skolor kommer elever med högre betyg som har lägre kunskapsnivåer att konkurrera ut elever som har högre kunskapsnivåer men som har haft lärare som är striktare i sin bedömning. Det riskerar att drabba den enskilda individen hårt. Även den elev som till följd av ovanligt lågt ställda krav från lärarens sida får ett godkänt betyg och kommer in på en utbildning riskerar att drabbas om hon eller han inte fått lära sig det som behövs för att klara utbildningen. Betyg som rättvisande speglar elevens kunskapsnivå ger både eleven och omgivningen en bättre bild och förståelse av denna kunskapsnivå, vilket i sin tur kan underlätta när personen i

fråga ska söka sig till en utbildning på högre nivå. På nationell nivå är det också viktigt att betygssystemet är rättssäkert, likvärdigt och stabilt. De betyg som sätts inom skolväsendet ska på ett rättvisande sätt spegla en persons kunskaper, inte t.ex. vilken skola hon eller han har gått på.

Betyg och meritvärden som inte upplevs som rättvisande och likvärdiga kan leda till sänkt tilltro och legitimitet för hela betygssystemet och skolans kunskapsuppdrag. Rättssäkerheten för eleverna och värdet av betygen som urvalsinstrument till högre studier urholkas. Bristande likvärdighet försvårar även uppföljningen av utbildningens kvalitet samt av elevers kunskapsnivåer över tid. Den kan också bidra till en utveckling där betygen stiger i förhållande till kunskapsnivån, dvs. leda till en betygsinflation. Det finns även en risk för otillbörlig påverkan på lärares betygssättning, t.ex. från skolledningen eller vårdnadshavare om betyg används som konkurrensmedel mellan skolor.

Betygsättningen är inte likvärdig

Det finns många rapporter som visar att elevers kunskapsnivåer inte alltid motsvarar de betyg som har satts och att betygssättningen varierar mellan, och ibland även inom, skolor och utbildningsanordnare. Betygsättningen är alltså inte alltid likvärdig. Det finns också utsagor om påtryckningar gentemot betygssättande lärare och om s.k. glädjebetyg. Den bristande likvärdigheten beskrivs bl.a. i Statens skolverks analyser av betygssättningen i grundskolan och gymnasieskolan (se bl.a. rapporterna *Analyser av likvärdig betygssättning mellan elevgrupper och skolor*, 2019, och *Analyser av likvärdig betygssättning i gymnasieskolan*, 2020) samt i SNS-rapporterna *Likvärdig kunskapsbedömning i och av den svenska skolan* (2014) och *Resultat och betygssättning i gymnasiefriskolor* (2022).

En förklaring till den bristande likvärdigheten är enligt Skolverket svårigheten för lärare på olika skolor att finna en samsyn i bedömningen. Betygsättningen är i stället i hög grad relativ på så sätt att den förhåller sig till den genomsnittliga prestationsnivån hos eleverna på skolan. En elev som lyckas väl på ett nationellt prov och som går på en skola där många andra elever också får höga provbetyg har lägre sannolikhet att få ett högt slutbetyg. I skolor med låga

resultat på det nationella provet råder det omvända förhållandet och det är lättare för eleven i fråga att få ett högt slutbetyg. Betygssättningen av elevernas prestationer sker utifrån en tolkning av betygs-kriterier och de progressionsuttryck som finns för att särskilja de olika betygsstegen. Det är enligt Skolverket en övermäktig uppgift för lärarna att åstadkomma en gemensam tolkning som leder till en nationellt likvärdig betygssättning av eleverna. De grundläggande problemen är inbyggda i själva betygssystemet och kan inte tillskrivas lärarkollektivet.

En annan aspekt som ibland framhålls är att systemet medger att det på skolor medvetet kan sättas för höga betyg, s.k. glädjebetyg. Att en skola har goda resultat i form av höga betyg lockar elever och vårdnadshavare. I ett system där skolor konkurrerar om elever och när god genomströmning är viktigt även av ekonomiska skäl kan incitament för betygsinflation skapas. Det behövs därför åtgärder som motverkar incitament att sätta högre betyg än vad som motiveras av elevers kunskapsnivå.

För att nå en markant ökad likvärdighet i betygssättningen krävs enligt Skolverket systemförändringar i betygssystemet, exempelvis att de nationella proven på något vis styr betygssättningen på skolnivå. Sådana systemförändringar behöver utredas noggrant och vägas mot tänkbara negativa effekter som förändringarna kan få (Likvärdiga betyg och meritvärden, Skolverket 2020).

I många länder spelar elevens betyg en mindre roll vid antagning till utbildning än i Sverige. Betyg kompletteras ofta med eller ersätts helt av resultat på olika typer av examensprov eller ämnesprov, åtminstone på gymnasial nivå. Det gäller bl.a. i våra nordiska grannländer.

Åtgärder för en mer likvärdig betygssättning är nödvändiga

Flera åtgärder har på senare år vidtagits för att främja en mer likvärdig betygssättning. Det pågår bl.a. förberedelser för central rättning av de nationella proven och Statens skolinspektion har i uppdrag att arbeta för en mer likvärdig bedömning och betygssättning. Trots detta kvarstår behovet av att se över betygssystemet, något som också lyfts av Riksrevisionen som vid sin granskning av statens insatser för likvärdighet i betygssättningen rekommenderat reger-

ingen att utreda hur betygssättningen kan bli mer likvärdig (Statens insatser för likvärdig betygssättning – skillnaden mellan betyg och resultat på nationella prov, RiR 2022:22). Enligt Riksrevisionen kan en utredning bl.a. undersöka möjligheten att formalisera acceptabla avvikelsemått på grupp- eller skolnivå, som kan användas i Skolinspektionens granskande verksamhet. Skolinspektionens tillsyn och kvalitetsgranskning utgår från de krav och mål som finns i skollagstiftningen och myndigheten är i behov av bättre förutsättningar för att kunna granska betygssättningen ur ett likvärdighetsperspektiv.

Skolverket har påbörjat ett arbete med att undersöka några olika modeller som skulle kunna användas för mer likvärdiga betyg och meritvärden. De olika modeller som Skolverket undersökt närmare innebär större systemförändringar och de utnyttjar alla på något sätt resultaten från ett nationellt prov, nationellt slutprov eller examensprov. Till de modeller som betraktas som potentiella system för mer rättvisande och likvärdig betygssättning hör examensprovmodellen, där centralt rättade examensprov används vid urval; direktstyrningsmodellen där läraren anpassar betygen efter skolans genomsnittliga resultat på nationella prov och sammanvägningsmodellen där elevens betyg vägs samman med elevens resultat på nationella prov till ett meritvärde som används för urval. Modereringsmodellen i den form som beskrivs i Skolverkets rapport bedöms vara mindre lämplig (se vidare nedan). Dessa modeller har dock inte analyserats ur alla aspekter och än så länge är de undersökta endast för grundskolan (Likvärdiga betyg och meritvärden, Skolverket 2020). Det finns därför skäl att med utgångspunkt i bl.a. Skolverkets rapport om mer likvärdiga betyg och meritvärden vidare analysera frågan om systemförändringar inom området. Vinster med sådana systemförändringar i form av förbättringar av betygssystemet och meritvärderingen behöver analyseras tillsammans med konsekvenser för skolans uppdrag att bidra till elevers kunskapsutveckling och för individen. Finns det t.ex. risker för att undervisningen snävas av mot det som är enkelt att mäta och att eleven därmed inte utvecklar alla de kunskaper eller förmågor som utbildningen ska syfta till, eller att vissa ämnen drabbas negativt av en föreslagen förändring? Samtidigt är problemen med likvärdigheten i betygssättning som har uppdragats så allvarliga att dessa riskerar att urholka skolans kunskapsuppdrag, vilket är ett starkt motiv för att genomföra åtgärder.

Ett väl utformat betygssystem bör harmoniera med skolans uppdrag att förmedla kunskaper och därmed främja elevernas kunskapsutveckling. Det är också viktigt att en förändring av systemet går att förstå, förklara och genomföra. Den modell som Skolverket förordar ur likvärdighetsaspekt är den s.k. modereringsmodellen. Enkelt kan den förklaras med att läraren sätter betyg oberoende av resultat på nationella prov och att detta betyg sedan modereras till ett betygsvärde efter hur skolans betyg förhåller sig till skolans resultat på nationella prov. Slutligen vägs detta modererade betygsvärde samman med elevens individuella resultat på det centralt bedömda nationella slutprovet. Även om modellen har fördelar är det en modell som är svår att förstå sig på och den omöjliggör för en elev att i förväg veta sitt meritvärde. Ett system kan framstå som effektivt för att skapa mer likvärdighet, men det behövs också legitimitet hos verksamheten, avnämare, lärare och elever i de förändringar som föreslås om de ska kunna fungera i praktiken. Mot denna bakgrund görs bedömningen att modereringsmodellen i den form som beskrivs i Skolverkets rapport är mindre lämplig. Det hindrar dock inte att även denna modell kan utgöra ett underlag i arbetet och att delar av modellen, eller en version av denna, kan ligga till grund för förslag inom området.

De modeller som Skolverket har undersökt närmare är alla utformade utifrån ämnen som har nationella prov. Betygsättningen i ämnen som inte har nationella prov varierar dock också mycket. I jämförelser mellan betyg i s.k. provämnerna, såsom matematik, svenska och engelska, och betyg i ämnen som inte har nationella prov, t.ex. bild och slöjd, är det flera skolor som har högre betyg i genomsnitt i de senare ämnena, dvs. i ämnen utan nationella prov. Att betygsättning varierar mellan ämnen kan i vissa fall ha sina förklaringar, men förekomsten av nationella prov förefaller ändå ha en viss återhållande effekt på de betyg som sätts. Det behöver därför också undersökas om det behövs systemförändringar när det gäller betygsättning i, och meritvärdering utifrån, ämnen som inte har nationella prov.

Det finns mot denna bakgrund starka skäl att se över hur systemet för betygsättning och meritvärdering i svensk skola kan bli mer rättvisande och likvärdigt.

Utredaren ska därför

- beskriva vilka system för betygssättning och meritvärdering som används i ett urval av nordiska länder och ett urval av andra med Sverige jämförbara länder,
- analysera för- och nackdelar med systemet för betygssättning och meritvärdering i det svenska skolväsendet i relation till andra länders system, till av Skolverket presenterade modeller och till av utredaren vidareutvecklade eller framtagna modeller,
- lämna ett huvudförslag, samt ett alternativt förslag på en systemförändring som innebär att resultat på nationella prov, examensprov, eller liknande, kan användas för att öka likvärdigheten i betyg eller meritvärden,
- lämna förslag på hur betygsinflation och bristande likvärdighet i betygssättning kan motverkas även i ämnen som inte har nationella prov,
- vid behov föreslå andra ändamålsenliga åtgärder som syftar till att främja mer likvärdiga betyg och meritvärden och motverka betygsinflation,
- beroende av övriga förslag ta ställning till behovet av förändringar av det nationella provsystemet och utbudet av nationella prov samt vid behov lämna förslag på sådana förändringar,
- analysera behovet av ändringar i angränsande frågor i relation till utredningens förslag, t.ex. när det gäller prövning, gymnasieexamen och behörighet till nationella program samt vid behov lämna förslag på sådana ändringar,
- analysera och beakta hur utformningen av olika system för betygssättning och meritvärdering kan påverka undervisningen, lärares arbetsituation, den administrativa bördan, genomströmning, samt olika elevgruppers motivation och välbefinnande samt möjlighet att få betyg och meritvärden som återspeglar deras kunskaper,
- beskriva hur de förslag som lämnas kan bidra till en ökad tilltro till systemet för betygssättning och meritvärdering hos lärare, elever och i samhället i övrigt,

- beskriva hur de förslag som lämnas kan bidra till att stärka Skolinspektionens möjligheter att granska skolors betygssättning, och
- lämna nödvändiga författningsförslag.

Översynen ska omfatta grundskolan, specialskolan, gymnasieskolan och kommunal vuxenutbildning på gymnasial nivå, dvs. skolformer där betygen används för tillträde till utbildning. Utredaren är dock oförhindrad att vid behov lämna förslag som rör andra skolformer.

Konsekvensbeskrivningar

Utredaren ska redogöra för ekonomiska och andra konsekvenser av sina förslag. Utöver vad som följer av kommittéförordningen (1998:1474) ska utredaren redogöra för förslagets förenlighet med Sveriges internationella åtaganden om mänskliga rättigheter inklusive principen om likabehandling samt särskilt ha ett barnrättsperspektiv i de analyser som görs och redovisa förslagets konsekvenser utifrån såväl FN:s konvention om barnets rättigheter som FN:s konvention om rättigheter för personer med funktionsnedsättning.

Konsekvensutredningen ska bland annat innehålla en analys av eventuella konsekvenser för lärare och lärares administrativa arbetsuppgifter. Utredaren ska även redovisa konsekvenser för elever, hur förslagen påverkar uppföljningen och utvärderingen av elevers kunskapsutveckling och konsekvenser för tillträde till utbildning på högre nivå.

De förslag utredningen lämnar ska ha sin grund i en samhälls-ekonomisk problemanalys. Vidare ska alternativa lösningar som övervägts beskrivas liksom skälen till att de har valts bort. Utredaren ska vidare göra en analys utifrån dataskyddssynpunkt av sådana behandlingar av personuppgifter som är en nödvändig konsekvens av utredningens förslag.

Kontakter och redovisning av uppdraget

I arbetet ska utredaren inhämta synpunkter från myndigheter och organisationer med relevans för uppdraget, bl.a. Statens skolverk, Statens skolinspektion, Specialpedagogiska skolmyndigheten, Statens

skolforskningsinstitut, Universitets- och högskolerådet, Sveriges Kommuner och Regioner, Friskolornas riksförbund, Idéburna skolors riksförbund, Sveriges skolledare, Sveriges lärare, Sveriges elevråd, Sveriges elevkårer och Elevernas riksförbund. Utredaren ska också hålla sig orienterad om andra utredningar inom kommittéväsendet med relevans för uppdraget.

Uppdraget ska redovisas senast den 21 februari 2025.

(Utbildningsdepartementet)

Betyg, bedömning och centrala prov i ett urval av länder samt International Baccalaureate Diploma Program (IB)

I denna bilaga beskrivs andra länders nuvarande samt International Baccalaureate Diploma Programs (IB) system för betyg, bedömning och centrala prov med fokus på aspekter som är relevanta för utredningens uppdrag. Fokus är på karakteristika för valda länders betygssystem, förekomst av centrala prov och deras syften samt betydelse för betyg. För länder där konstruktion och genomförande av centrala prov har visat sig särskilt intressanta i relation till de modeller som utredningen ska undersöka redovisas fördjupad information om dessa. I slutet av bilagan finns översiktstabeller över undersökta aspekter för länderna.

Uppgifterna har sammanställts huvudsakligen baserat på egna kontakter med olika länders utbildningsmyndigheter men även med grund i aktuella internationella sammanställningar och publicerad information.

Genomgången är huvudsakligen genomförd mellan oktober 2023 och juli 2024 och läsaren bör vara medveten om att delar av informationen riskerar att bli inaktuell eftersom utbildningssystem reformeras med jämna och ojämna mellanrum.¹

¹ Översiktliga, breda jämförelser mellan nordiska och europeiska länders utbildningssystem har tidigare gjorts i bland annat betänkandet *Bygga, bedöma, betygssätta – betyg som bättre motsvarar elevernas kunskaper* (SOU 2020:43) och *Tio EU-länders utbildningssystem: En översiktlig bild* (Skolverket 2007). I rapporten *Betygssystem i internationell belysning* (Skolverket 2016) finns också ett flertal länders betygssystem beskrivna. Även för dessa rapporter gäller att information i dem kan vara inaktuell efter genomförda reformer.

Urval av aspekter för beskrivningen av olika länders system

Följande har identifierats som relevanta aspekter att beskriva fördjupat för ett urval av länder:

1. Grundläggande om betygssystemet
 - Är betygen standardrelaterade, målrelaterade eller normrelativa?²
 - Hur är betygsskalan konstruerad och hur många betygssteg finns det?
 - Vilka målnivåer finns och hur ser betygs-kriterier/*standards* eller motsvarande ut?
 - Från vilken ålder sätts betyg?
 - Används betygen för behörighet och/eller urval om ett sådant sker efter grundskolan eller efter gymnasieskolan?
 - Går betyg att överklaga?
2. Används centrala prov och/eller examensprov?³
 - Om centrala prov används och i så fall i vilket syfte.
 - Om examensprovsresultat används, och i så fall hur, för antagning och behörighet efter grundskolans respektive gymnasieskolans slut samt om de används ensamma som grund för antagning och urval eller tillsammans med andra antagningsinstrument.
 - Hur konstrueras proven?
 - Hur hanteras svårbedömda uppgiftsformat, t.ex. muntliga prov?
 - Är proven jämförbara över tid?
 - Hur används proven i skolan?
 - Hur genomförs proven (av externa examinatorer/egna lärare, digitalt/analogt)?

² Målrelaterad bedömning är bedömning utifrån läroplanens mål. En kriteriebaserad bedömning uttrycker karakteristika eller kännetecken på kunskap. Standardbaserad bedömning är en specificerad variant med bestämda kvalitetsnivåer för respektive betygssteg. Jfr Sadler (1987), refererad i *Betygssystem i internationell belysning* (Skolverket 2016, s. 7).

³ Med centrala prov avser vi prov som konstrueras, kvalitetssäkras och distribueras från nationell eller regional nivå. Examensprov avser prov som krävs för att få en viss examen. De senare kan vara centralt framtagna men det förekommer också att de konstrueras lokalt.

- Vem bedömer proven (egna lärare/externa bedömare/automatik)?
3. Andra urvalsgrunder som används för ansökan till högre utbildning än examensprov/betyg.

Danmark⁴

Skol- och betygssystem

I Danmark ges betyg fr.o.m. årskurs 8 i den nioåriga grundskolan. En tionde årskurs erbjuds för den som önskar och behöver det. Därefter följer en vanligtvis treårig gymnasial utbildning.

Danmark har en sjugradig betygsskala som går från -3 , som är det lägsta betyget, till 12 , som är det högsta betyget. Skalan bygger på ECTS-skalans sju steg med två underkända betygssteg F och Fx, som i det danska systemet motsvaras av -3 respektive 00 .⁵

De lärarsatta betygen sätts målrelaterat utifrån målen för ämnet och utbildningen. Betygsättningen i relation till målen är kompensatorisk, där styrkor kan vägas mot svagheter. I riket ska betygen förhålla sig till en relativ fördelning för de godkända betygsstegen. Den tänkta fördelningen av de godkända eleverna i riket innebär att de bästa 10 procenten ges A, de följande 25 B, de mellersta 30 C, därefter 25 D och de svagaste 10 procenten ges E.⁶

Folkeskolens Nationale Overgangstest

I kvalitetssäkrande syfte finns sedan år 2022 nationella prov i läsning och engelska i flera av grundskolans årskurser. Dessa kallas sammanfattande för *Folkeskolens Nationale Overgangstest*.⁷ Från och med 2026 kommer proven att ersättas med *Folkeskolens Nationale Færdighedstest*, en ny form av digitala, automaträttade prov i matematik och läsning.

Proven ska genomföras i början av läsåret så att de kan bidra till lärarnas organisering av undervisningen under resten av läsåret. Proven distribueras digitalt och eleverna genomför de skriftliga proven direkt

⁴ Huvudsaklig källa för avsnittet: det danska Børne- og Undervisningsministeriets hemsida, <https://www.uvm.dk>.

⁵ A = 12, B = 10, C = 7, D = 4, E = 02, Fx = 00, F = -3 .

⁶ SOU 2020:43, s. 330–333.

⁷ Børne- og Undervisningsministeriet (2024).

i ett digitalt prov- och examinationssystem. Rektor kan på grundval av en individuell bedömning erbjuda särskilda villkor för elever med psykiska eller fysiska funktionshinder eller liknande svårigheter samt för tvåspråkiga elever med bristande kunskaper i danska, när detta är nödvändigt för att dessa elever ska vara jämställda med andra i provsituationen.⁸ Det kan t.ex. handla om hjälpmedel, tilläggstid eller medhjälpare men ska ske i samma omfattning och form som eleven får stöd för att klara sitt ordinarie skolarbete. Rektor kan även bevilja dispens från att skriva proven om eleven bedöms ha betydande funktionsnedsättning eller bristande kunskaper i danska, men det krävs då att eleven och dess vårdnadshavare är eniga med rektor.

Det finns även nationellt framtagna digitala prov i engelska, fysik/kemi, biologi, geografi och danska som andraspråk som skolor på frivillig basis kan välja att använda till och med läsåret 2025/26. Danska grundskolor med årskurs 8, 9 eller 10 kan för närvarande delta i en tvåårig försöksverksamhet med nya betygs- och bedömningsformer i grundskolan.⁹ Skolorna kan välja mellan tre alternativa försöksverksamheter: sätta betyg för arbetsprestation, sätta betyg i ordning och uppförande eller skriftlig formativ återkoppling och bedömning i stället för betyg.

Examensprov i grundskolan och gymnasieskolan

Examensproven i grundskolan konstrueras av särskilda uppgifts- och testkommissioner på uppdrag av Danmarks utbildningsdepartement. Medlemmarna i kommissionerna ska ha praktisk erfarenhet från undervisning och/eller forskning i ämnet. Skolans rektor utfärdar intyg för grundskoleexamen till elever i årskurs 9 som har uppnått ett visst genomsnitt på obligatoriska prov i muntlig danska, skriftlig danska, matematik, naturvetenskap, engelska och två ämnen som varierar mellan skolor. Ett valfritt prov i ett praktiskt ämne i årskurs 8 kan också ingå. Proven, med undantag av muntliga prov som hanteras lokalt, bedöms externt och bedömningen kan överklagas. Intyget ska utöver provresultaten innehålla information om lärarsatta betyg i övriga ämnen.¹⁰

⁸ Förordning: Bekendtgørelse om obligatoriske test i folkeskolen i skoleårene 2022/23 til og med 2025/26 (Folkeskolens Nationale Overgangstest).

⁹ Børne- og Undervisningsministeriet (2023a).

¹⁰ Børne- og Undervisningsministeriet (2023b).

I princip alla som vill gå en yrkesutbildning på gymnasial nivå, *erhvervsuddannelse*, antas till sådan så länge man har godkända grundskolebetyg i danska och matematik. Eleverna ansvarar dock själva för att få en lärlingsplats för praktik, vilken vanligen berättigar till studielön.¹¹ En särskild form av teori-kombinerad gymnasial yrkesutbildning, *eux*, kan ge behörighet till vissa högskolestudier.¹²

Det krävs godkänd grundskoleexamen, oftast med minst 5,0 i genomsnittsbetyg i både lärarsatta betyg och på examensproven, för att vara behörig till högskoleförberedande gymnasiestudier. Antagningen görs baserat på betygsgenomsnittet med vissa möjliga undantag och krav på lägsta betyg samt med hänsyn till restid till sökt skola.¹³

I den högskoleförberedande gymnasieskolan sätts terminsbetyg i alla ämnen.¹⁴ För att få gymnasieexamen krävs också examensprov i minst tio ämnen. Proven kan vara skriftliga, muntliga, i projektform eller ha blandade bedömningsformer. Vissa av proven kan genomföras efter årskurs 1 och 2, medan slutprov i fördjupande ämnen och danska genomförs efter år tre. De skriftliga proven, som i huvudsak görs digitalt, bedöms av två externa examinatorer. Muntliga provdelar bedöms av elevens egen lärare och en extern censor. I exempelrutan nedan framgår mer detaljerat hur upplägget ser ut ur elevens perspektiv.

Gymnasiestudierna avslutas med att ett examensbevis som innehåller flera olika betyg utfärdas: lärarsatta betyg, examensprovsbetygen, betyg på projektarbete, viktade examensbetyg och genomsnittsbetyget.¹⁵ Det sistnämnda används för urval vid ansökan till högre utbildning. Om en elev saknar ett examensprovresultat blir den posten i avgångsbetyget tom. Det ges nya tillfällen att skriva examensprov i augusti, men detta sker efter att ansökan till högskolan för höstterminen ska vara inlämnad. Tidigare uppges vissa universitet ha gett dispens med att komplettera betyg, men ökad konkurrens innebär att de inte längre anser sig behöva göra detta. Eleven får i stället vänta med sin ansökan till framtida tillfällen.

Skolor anmäler själva till Styrelsen för Undervisning och Kvalitet (under Børne- og Undervisningsministeriet) vilka av deras lärare som

¹¹ Øresunddirekt (2024).

¹² UddannelsesGuiden (2024a).

¹³ UddannelsesGuiden (2024b).

¹⁴ Det finns sex huvudsakliga valmöjligheter på inriktning: naturvetenskap, samhällsvetenskap, språk eller konst inom generell gymnasieexamen (STX) samt ekonomisk respektive teknisk examen.

¹⁵ SOU 2020:43, s. 330–333 och Eurydice (2024a).

kan vara externa censorer och i vilka ämnen de är behöriga. Det är obligatoriskt att delta i muntlig examination men det krävs inte att man registrerar sig som skriftlig censor. På grundval av anmälningarna väljer och fördelar Styrelsen for Undervisning og Kvalitet de externa censorerna. Principerna för fördelning är geografisk spridning, kön och tidigare erfarenhet. En lärare får dock inte vara extern censor hos samma skola två år i rad.

I Danmark används två kvotgrupper för urval till högskolan och lärosätena bestämmer själva gruppernas storlek med förbehållet att samtliga platser inte får fördelas med betyg (*kvote 1*) som urvalsgrund. Antagningen i *kvote 2* ska grundas på av lärosätena valda och specificerade urvalskriterier där t.ex. motivation, arbetslivserfarenhet, ideellt arbete och utlandsvistelse kan vägas in men det förekommer även intervjuer, antagningsprov och personliga brev som urvalsgrunder.¹⁶

Exemplet Danmark:¹⁷ Undervisningen för danska elever på högskoleförberedande gymnasieutbildningar (STX) avslutas i mitten av maj och då ska lärarna ha satt sina betyg. Ingen undervisning får ske efter att betygen har satts i maj, utöver viss repetition inför examensprov.

I flertalet ämnen, även matematik och NO-ämnen, får eleverna ett betyg på muntlig och ett på skriftlig förmåga. Det är ett krav att gymnasieskolorna ska låta eleverna genomföra fem större prov under utbildningen, s.k. *årsprøver*. Tre av dessa ska vara muntliga eller skriftliga i danska, matematik och ett ämne som är karakteristiskt för vald studieinriktning. Skolan väljer själv övriga två ämnen. Dessa prov konstrueras lokalt, eller så använder lärarna gamla examensprov, och de ses som viktiga för att eleverna ska träna sig på provsituationer inför examensproven. Löpande terminsprov och andra examinationer bestämmer skolorna själva över och ansvarar för bedömningen av.

Den s.k. nationella provplanen, som visar vilka provämnen samt vilka prov som kommer att vara muntliga respektive skriftliga, offentliggörs först i maj så att *teach-to-the-test*-effekterna ska minimeras i undervisningen.

¹⁶ UddannelsesGuiden (2024c).

¹⁷ Styrelsen for Undervisning og Kvalitet (2024), Studienet.dk (u.å.).

I slutet av första året genomför varje elev ett prov i allmän språkförståelse och ett i grundläggande naturkunskap. Dessa räknas inte till studentexamen. Vanligtvis gör eleverna dock ytterligare minst ett muntligt eller skriftligt examensprov i ett ämne som då avslutas. Betygen i dessa examensämnen, både de lärarsatta betygen och provbetygen, förs in i examensbeviset vid studentexamen.

Andra gymnasieåret avslutas i mitten av maj på liknande sätt och eleverna genomför examensprov i ett till tre ämnen som då avslutas. I slutet av tredje och sista året genomför eleverna resterande examensprov. Man kan normalt sett göra ett prov per dag och provtiden per ämne är fyra till sex timmar. För en typisk sistaårselev innebär det fem till åtta dagar med prov under tre till fyra veckor.

Resultaten delges eleverna efter rättning i juni. Bedömningen ska vara klar senast till sista ansökningsdag till högskolan, normalt sett runt den 5 juli.

Elevexempel:

Lisa läser allmän studentexamen, STX, med naturvetenskaplig inriktning. Första året skriver hon de obligatoriska proven i allmän språkförståelse och grundläggande naturkunskap samt gör ett muntligt examensprov i franska. Andra året skriver hon examensprov i historia och informationsteknik samt gör ett muntligt examensprov i kemi. Sista året gör Lisa skriftliga examensprov i danska, engelska, matematik och antikens studier samt muntliga examensprov i danska och matematik. Hon får också göra en muntlig presentation av sitt skriftliga specialarbete som betygsätts som ett examensprov.

Norge¹⁸

Skol- och betygssystem

Norges tioåriga grundskola påbörjas vid sex års ålder och har en betygsskala från 1 till 6 där 1 är det lägsta och 6 är det högsta betyget. Betyg sätts från och med årskurs 8, dvs. motsvarande årskurs 7 i

¹⁸ Huvudsaklig källa för avsnittet: det norska Utdanningsdirektoratets hemsida, <https://www.udir.no>.

Sverige.¹⁹ Betygssystemet är målrelaterat och elevens kunskaper bedöms utifrån läroplanens kompetensmål.²⁰ Det finns skrivna betygskriterier, *kjennetegn*, för tre av betygsstegen: betyget 2 som står för låg kompetens i ämnet, 4 god kompetens i ämnet och 6 utmärkt kompetens i ämnet.²¹ I grundskolan räknas även betyget 1 som ett godkänt betyg, medan det i gymnasieskolan räknas som underkänt; där är betygen 2–6 godkända.²² Betygen betraktas som provisoriska, *underveisvurdering*, fram till grundskolans sista och tionde årskurs. Betyg sätts i ordning och uppförande, vanligen med vitsord i stället för siffror, och dessa tillmäts inte vikt vid ansökan till fortsatta studier.

Gymnasieutbildningen är uppdelad i en allmän gymnasieskola och yrkesutbildning. Urval och antagning bestäms på regional nivå och principerna varierar. Låga grundskolebetyg är inget hinder för antagning eftersom det inte finns någon godkäntgräns. Yrkesutbildning ger normalt sett inte behörighet till högskola men det finns möjlighet att få det med ett års kompletterande studier.

Den allmänna gymnasieskolan är treårig och ger behörighet till högre utbildning.²³ Både lärarsatta betyg i respektive ämne utifrån en bedömning av olika betygsunderlag och elevens betyg på examensproven anges i avgångsbetyget.²⁴ I språkämnerna skiljer man på skriftlig och muntlig framställning och en elev får därför totalt fyra betyg i t.ex. norska respektive engelska.

Nationella prov och examensprov

I årskurserna 5, 8 och 9 görs nationella prov i norska, matematik och engelska i syfte att mäta kunskaper på individ-, grupp-, skol- och lokal/regional nivå.²⁵ Proven tas fram i samarbete med lärosäten och har ankaruppgifter för att möjliggöra jämförelser över tid. Resultaten ska användas som underlag för kvalitetsutveckling.

Grundskolan avslutas med examensprov i ett av ämnena norska, samiska, engelska eller matematik. Vilket ämne som en viss elevgrupp ges provet i bestäms lokalt av skolan och meddelas eleverna

¹⁹ Eurydice (2024b).

²⁰ SOU 2020:43, s. 333.

²¹ Utdanningsdirektoratet (2020).

²² SOU 2020:43, s. 334.

²³ Eurydice (2024a).

²⁴ Eurydice (2024b).

²⁵ Utdanningsdirektoratet (2023).

ett par dagar innan provet genomförs.²⁶ De skriftliga proven konstrueras centralt av nämnder som består av lärare och andra yrkesverksamma, totalt tre eller fyra medlemmar i varje nämnd. Deras arbete samordnas av ämnesspecialister från Utdanningsdirektoratet och ska bygga på framtagna kvalitetskriterier för utvecklingen av prov.²⁷ Utvärderingar visar dock att det finns problem med att proven inte är stabila över tid.²⁸ Från och med år 2023 ska så gott som alla elever skriva proven digitalt men det förekommer fortfarande vissa prov tryckta på papper. Proven bedöms centralt, dvs. inte av de egna lärarna.

Utöver de centralt konstruerade proven förekommer regionalt och lokalt konstruerade prov som examineras av censorer. Examinatorerna i de olika ämnena föreslås av rektorer och utses sen på regional nivå. Examinatorerna ska vara ämneslärare i ämnet och de får inte examinera på en skola där de undervisar. Alla uppgifter bedöms av två examinatorer. När de sätter betyg har de inte tillgång till den andra examinatorns betygsförslag. Examinatorsutbildningar hålls regionalt och leds av handledare som utbildats av Utdanningsdirektoratet. Syftet med utbildningen är att trygga bedömaröverensstämmelsen. En examinator får cirka 150 prov som ska bedömas under en treveckorsperiod. När alla prov har bedömts av två examinatorer träffas alla för en gemensam bedömning där det slutliga betyget fastställs. Det är först under det gemensamma bedömningstillfället som examinatorerna har tillgång till andra censorers föreslagna betyg. Bedömningen sker i ett digitalt system som utvecklats för test- och provadministration. Examensprovsbetygen blir en del av elevens avgångsbetyg från grundskolan men utgör endast cirka 20 % av vikten i betyget.²⁹

Huvuddelen av de skriftliga proven är digitala och administreras centralt på liknande sätt som proven i grundskolan.³⁰ Eleverna genomför examensprov i minst tre ämnen. Proven som tas fram centralt prövas delvis ut men de har inga ankaruppgifter. Försök med ankaruppgifter i matematik har gjorts och visat på små variationer mellan år. Det förekommer också praktiska och muntliga prov som tas fram och arrangeras lokalt eller regionalt. Bedömningen av både de skriftliga

²⁶ SOU 2020:43, s. 334 och Eurydice (2024c).

²⁷ Utdanningsdirektoratet (2021).

²⁸ Björnsson (2020).

²⁹ Utdanningsdirektoratet (2022).

³⁰ Utdanningsdirektoratet (2024).

och muntliga delarna görs externt av censorer.³¹ Censorerna samråder och enas om hur bedömningen ska göras i respektive ämne i syfte att eftersträva likvärdig bedömning. I slutbetyget skrivs både lärarens betyg och examensprovsresultatet in. Om eleven saknar examensprovsresultat blir den posten tom. Den får alltså ett betyg mindre än den normalt skulle ha haft. Det finns inget uppsamlingstillfälle i nära anslutning till ordinarie provtillfälle. Den som missar ett provtillfälle eller inte blir godkänd får göra om provet vid nästa examensperiod. Det finns två examensperioder per år, en på våren och en på hösten. Elever kan senare välja att skriva prov som privatister mot en avgift.

Meritpoäng kan spela stor roll vid antagning till eftergymnasial utbildning

Ansökan till eftergymnasial yrkesutbildning och högskoleutbildning görs via en gemensam webbportal, samordnaopptak.no. Antagningen till högskolan är i flera avseenden lik den svenska med allmän och särskild behörighet samt olika kvotgrupper.³² Hälften av platserna viks vanligen för sökande under 21 år som framför allt konkurrerar med sina gymnasiebetyg. En allmän kvotgrupp finns och där ges sökande meritpoäng baserat på ålder, betyg i ämnen som tagits efter gymnasieskolan samt högre betyg i upplästa ämnen. Även militärtjänst, yrkesutbildning, tidigare högskolestudier och folkhögskolestudier kan ge extra meritpoäng.³³ Grunden för meritvärdesberäkningen är att studentens genomsnittsbetyg från gymnasieskolan multipliceras med tio för att ge ett meritvärde mellan 20 och 60. Till det läggs meritpoäng för gymnasiestudier i STEM-ämnen, språk och, i förekommande fall, resultat på antagningsprov. Vid ansökan till utbildningar där könsfördelningen är ojämn får män respektive kvinnor extra meritpoäng. Det förekommer även vissa utbildningar som helt och hållet antar på basis av antagningsprov.

Antagningen till yrkeshögskolans utbildningar, som är 0,5–2 år långa, sker på liknande sätt som till högskolan. Först multipliceras genomsnittsbetyget från gymnasiet med tio för att ge ett meritvärde mellan 20 och 60 och sen läggs extra meritpoäng till för yrkes-

³¹ SOU 2020:43, s. 334.

³² Samordna opptak (2024a).

³³ Samordna opptak (2024b).

erfarenhet, tagna yrkes-/gesällbrev och särskilt goda vitsord i dessa samt i förekommande fall resultat på antagningsprov.³⁴

Finland³⁵

Skol och betygssystem

I Finland kan betyg sättas från och med den tioåriga grundskolans årskurs 1, men skolorna väljer själva om de vill sätta betyg i de tre lägsta årskurserna eller om elevernas kunskapsutveckling ska dokumenteras på annat sätt.³⁶ Betyg kan överklagas. Från och med årskurs 4 ska sifferbetyg sättas som komplement till eventuella andra omdömen. Uppförandebetyg i relation till den lokala läroplanen sätts varje termin men skrivs inte med i slutbetyget.

Finland har en sjugradig betygsskala som går från 4 till 10, där 4 är underkänd prestation och 5–10 är godkända betyg. Det finns skrivna kunskapskrav för vitsorden 5, 7, 8 och 9 som betyg sätts utifrån. Därutöver anges mål för undervisningen, innehåll, mål för lärandet och föremål för bedömning.³⁷ Ett avgångsbetyg från grundskolan sätts om eleven har fått godkända betyg i alla ämnen. I grundskolan finns det inga nationella prov med koppling till bedömning av enskilda elever.³⁸ Däremot förekommer urvalsbaseade prov för utvärdering av kunskapsutveckling i årskurserna 3, 6 och 9. Resultaten återkopplas på skolnivå i relation till hela landets resultat. Bedömningen görs av lärare och central stickprovskontroll görs för att utvärdera bedömaröverensstämmelsen men inga centrala justeringar görs eftersom det inte finns koppling till elevers betyg.

Den högskoleförberedande gymnasieskolan är normalt sett treårig men kan även genomföras på två eller fyra år för den som önskar ha en annan studietakt.³⁹ Lärarnas betygssättning i gymnasieskolan är löst reglerad. I de nationella grunderna för läroplanen överlämnas många frågor till den lokala nivån att besluta om, exempelvis betygens utformning.⁴⁰

³⁴ Samordna opptak (2024c).

³⁵ Huvudsakliga källor för avsnittet: Utbildningsstyrelsens och Studentexamensnämndens hemsidor, <https://www.oph.fi/sv> och <https://www.ylioppilastutkinto.fi/sv>.

³⁶ Utbildningsstyrelsen (2020).

³⁷ Utbildningsstyrelsen (2024a).

³⁸ SOU 2020:43, s. 335–339.

³⁹ Utbildningsstyrelsen (2024b).

⁴⁰ SOU 2020:43, s. 337 och Utbildningsstyrelsen (2019).

Studentexamen

Den finska studentexamen har en lång historia och bestod tidigare, precis som den svenska, av både skriftliga och muntliga examinationer. De muntliga examinationerna har avskaffats, dock har Finland behållit studentexamen.⁴¹ Studentexamen administreras av Studentexamensnämnden. Examensproven, som är helt och hållet digitala, tas fram nationellt och syftar till att pröva huruvida studenterna har tillägnat sig de kunskaper och färdigheter som föreskrivs i läroplanen och om eleverna har uppnått tillräckliga kunskaper enligt läroplansmålen.⁴²

Proven provas inte ut på förhand, utan man förlitar sig på att de konstrueras på ett tillförlitligt sätt. Diskussion förs om att eventuellt förbättra konstruktäckningen genom att införa ytterligare provdelar. Medvetenheten är stor om den starka styrning som examensproven har. Dock finns muntliga delar och praktiska moment såsom laborationer med som krav i lärarens betygssättning, vilket till viss del kan motverka att sådana delar faller bort även om de inte provas i studentexamensproven.

Studentexamensnämnden menar att det inte förs någon egentlig diskussion i landet kring ankaruppgifter eller utprovning av provfrågor på förhand. En aspekt som framhålls är snarare att proven anses så krävande i sig att det skulle vara oetiskt att utsätta elever för ännu fler prov/frågor än de som de faktiskt bedöms på.

Det finns tänkta fördelningar för provresultaten, men det är inte en fast fördelning för varje enskilt prov. Fördelningen som ska eftersträvas för de olika betygen/vitsorden är följande:

- 5 % laudatur (berömlig)
- 15 % eximia cum laude approbatur (med utmärkt beröm godkänd)
- 20 % magna cum laude approbatur (med beröm godkänd)
- 20 % cum laude approbatur (icke utan beröm godkänd)
- 20 % lubenter approbatur (godkänd)
- 15 % approbatur (icke fullt godkänd)
- 5 % improbatur (icke godkänd)⁴³

⁴¹ Petersson (2010, s. 53).

⁴² Studentexamensnämnden (2023).

⁴³ Studentexamensnämnden (2024a).

Preliminära bedömningsinstruktioner tas fram inför provgenomförandet. Dessa justeras efter att provet är genomfört beroende på utfall, t.ex. om det visade sig att många elever missförstod en viss provuppgift. Kravgränserna sätts sedan i efterhand med MST-modellen som innebär en typ av årsvis standardisering, men där standardiseringen görs för samtliga prov och inte varje enskilt prov var för sig. Det innebär att varje enskilt prov inte antar en specifik fördelning, utan att fördelningen läggs ut på de olika proven utifrån hur svåra de blev. Modellen har inte kritiserats särskilt mycket sedan de frångick fasta fördelningar (normalfördelning) för varje enskilt prov och övergick till MST-modellen.

Bedömning av examensproven görs först av lärare med stöd av kriterier och utifrån Studentexamensnämndens föreskrifter. Dessa bedömningar är preliminära. Därefter slutbedömer Studentexamensnämndens censorer alla prov.⁴⁴

Den som avlägger studentexamen ska skriva prov i fem ämnen och det är obligatoriskt med prov i modersmål och litteratur.⁴⁵ Övriga prov ska väljas från minst tre av följande grupper: matematik, det andra inhemska språket, främmande språk och realämne (t.ex. psykologi, filosofi, historia, fysik och biologi).⁴⁶ Den som så önskar kan avlägga extra prov utöver de fem prov som krävs för examen. Proven ges vår och höst och kan tas vid tre efterföljande provtillfällen och ändå räknas in i en och samma examen. Det går att kontrollera att ens tänkta kombination av provämnen uppfyller kraven i ett särskilt webbaserat verktyg, ILMO.⁴⁷

Samtliga 42 prov, varav runt hälften är i språk, finns i en finskspråkig och en svenskspråkig variant. Det finns också specialprov för personer med syn- respektive hörselnedsättning. Proven hålls hemliga tills de skrivs, därefter blir de offentliga. Rektor ansvarar för bedömning av och specialarrangemang för elever som bedöms behöva förlängd provtid. Man får skriva om examensprov ett obegränsat antal gånger, inklusive för att få bättre vitsord, och det går att bygga på med ytterligare prov i ämnen efter examen. Proven ges vår och höst och samtliga prov erbjuds alltid. Den som missar ett provtillfälle får vänta till nästa tillfälle.

⁴⁴ Studentexamensnämnden (2024a).

⁴⁵ Studentexamensnämnden (2024b).

⁴⁶ Studentexamensnämnden (2024c).

⁴⁷ ILMO, <https://ilmo.ylioppilastutkinto.fi/sv>.

Många prov har få provtagare. Eftersom det blir selekterade urval i ämnen med få elever, såsom avancerad fysik och ryska, hanteras betygssättningen i dessa manuellt. Antagandet är dock att provskrivarkohorter i respektive ämne är jämförbara över tid. Modellen för betygssättning utgår därför i huvudsak från medelvärden och fördelning av totala resultat. Inget betygsgenomsnitt från en enskild elevs samtliga prov räknas ut.

Yrkesutbildning och högskoleutbildning

Grundläggande yrkesutbildning på gymnasial nivå är vanligen också treårig men ger inte grundläggande behörighet till högskoleutbildning om man inte själv väljer att komplettera med extra kurser och examensprov. Däremot kan man söka till yrkeshögskolan, vilken i Finland ger vissa utbildningar som i Sverige är högskoleutbildningar, t.ex. barnmorska, fysioterapeut, sjuksköterska och vissa ingenjörsutbildningar.⁴⁸

Finska lärosäten har stor autonomi och antar på egna beslutade behörighetskrav. Ofta är studentexamen ett grundläggande behörighetskrav för tillträde. Lärosäten koordinerar sig själva när det gäller behov av behörighetskrav och urvalsgrunder till likartade utbildningar. Urvalsprov liknande det svenska högskoleprovet finns men det förekommer även lokala lärosäteskonstruerade antagningsprov.

England⁴⁹

Skol- och betygssystem

England har skolplikt mellan fem och sexton års ålder. Därefter tar utbildningsplikt vid fram till 18 års ålder. Första delen av grundskolan, *primary school*, börjar vid fem års ålder och varar till och med 11 års ålder.⁵⁰ Inga omdömen eller betyg används för antagning till fortsatt utbildning utan alla kan gå vidare till nivån *secondary school* i en så kallad *comprehensive school*, dvs. en offentligt finansierad skola. *Upper-secondary school*, den tredje och sista delen av grund-

⁴⁸ infofinland.fi (2024).

⁴⁹ Huvudsaklig källa för avsnittet: muntliga uppgifter från Cambridge Assessment International.

⁵⁰ Standards and Testing Agency (2024).

skolan, och privatfinansierade skolor ställer däremot ibland upp specificerade antagningskrav och/eller har antagningsprov.

Under tiden i grundskolan kan lärare ge eleverna omdömen kontinuerligt men de får inga formella lärarsatta betyg. Skolor kan i huvudsak välja skala för betyg och omdömen själva. I skolår 10 och 11, dvs. den sista delen av grundskolan, är det vanligt att lärarna använder samma niogradiga betygsskala som används för slutbetygen på examensproven för sin betygssättning. Det vänjer eleverna vid skalan inför slutproven.

Formativa prov i lägre årskurser

Eleverna skriver centralt framtagna nationella prov (*national curriculum tests*) i årskurserna 2 och 6. Proven är av formativ karaktär i syfte att bedöma huruvida elevernas kunskaper är i linje med det förväntade och identifiera behov av stöd. I årskurs 2 görs prov i engelska och matematik. Provbetyg ges på en skala mellan 85 och 115 där poäng under 100 indikerar behov av stöd (på väg mot förväntad nivå), 100 är förväntad nivå och över 100 över förväntad nivå. I årskurs 6 görs, förutom prov i samma ämnen som i årskurs 2, lärarbedömningar av elevernas läsförmåga och NO-kunskaper. Samma provbetygsskala och nivåbedömning används men med skillnaden att maxbetyget är 120. I NO ges ett av två omdömen: under förväntad nivå respektive på förväntad nivå.

High stakes-prov i grundskolan och på gymnasial nivå

11+-prov

I årskurs 6 kan eleverna skriva ett frivilligt så kallat 11+-prov som används för antagning till någon av landets 160 *grammar schools*. Avsikten med provet är att de bäst lämpade ska få en plats på skolorna. Systemet har dock kritiserats för att vara olikvärdigt av bl.a. tre skäl: ojämn geografisk spridning av skolenheter över landet,⁵¹ underrepresentation av elever från resurssvaga/socioekonomiskt utsatta

⁵¹ Johnes (2016).

familjer som skriver provet samt hög förekomst av privatundervisning/läxhjälp inför provet bland de elever som antas.⁵²

GCSE-prov

I slutet av årskurs 11 gör eleverna s.k. GCSE-prov (*General Certificate of Secondary Education*) i mellan nio och tolv ämnen. Proven bedöms och betygsätts externt. I engelska och matematik erbjuds provtillfällen två gånger om året. I övriga ämnen kan prov endast skrivas på våren. Något betygsmedelvärde eller liknande räknas inte ut. Den som får underkänt får göra om proven ett obegränsat antal gånger. Betygen kan användas som grund för antagning till gymnasieskola, *upper secondary school* och även till högskoleutbildning. Antagningen är dock inte centraliserad utan skolor och lärosäten bestämmer själva vad som krävs för behörighet och betyg, vilket innebär att det i praktiken går att söka till såväl gymnasial utbildning som högskola utan att ha gjort några examensprov. Vanligast är att det krävs särskilda förberedande kurser eller godkända antagningsprov för att få en plats om man saknar GCSE/A-levels.

Den ordinarie högskoleförberedande gymnasieskolan kräver normalt sett att man som lägst har betyget 4 eller 5 på minst fem GCSE-prov.

A-levels

Utbildningen på gymnasial nivå är tvåårig och eleverna läser vanligen tre eller fyra ämnen. I slutet av andra året skriver eleverna prov, *A-levels*, i sina ämnen. Fem olika *exam boards* erbjuder *A-levels*-prov och den enskilda skolan väljer själv samarbetspartner. Vissa universitet föredrar eller kräver dock att studenterna skrivit sina *A-levels*-prov med en viss *exam board*. Proven bedöms centralt och provresultaten konverteras till bokstavsbedyg på skalan E–A*.

A-levels-prov kan göras två till tre gånger om året. För den som missar något enstaka delprov med giltig anledning kan en *exam board* undantagsvis beräkna ett förväntat provresultat baserat på ett eller en kombination av följande så att eleven får ett godkänt betyg i ämnet:

⁵² Kirby (2016).

- a) elevens övriga genomförda delprov,
- b) elevens uppmätta prestationer under kursen,
- c) elevens resultat på eventuella träningsprov samt
- d) medelprestationen nationellt på det berörda delprovet.⁵³

Det förekommer också att eleven ges möjlighet att skriva provet i efterhand. Den som missar alla delprov i ett ämne får återkomma vid nästa provtillfälle.

Det förekommer varken ankaruppgifter i eller utprovning av proven. Proven är i grunden avgiftsbelagda men blir i praktiken vanligen avgiftsfria eller subventionerade för individen. Diskussioner om vilka *exam boards* som har enklast respektive svårast prov är vanligt förekommande bland elever och skolor.⁵⁴ Provkonstruktörerna eftersträvar en stabil svårighetsgrad hos proven över tid, men för att justera för fluktuationer görs den slutliga kravgränssättningen i efterhand.⁵⁵

Antagning till högskoleutbildning och yrkesutbildning

Bokstavsbetygen motsvarar i sin tur olika s.k. *UCAS-poäng* som ensamt eller tillsammans med andra underlag används för lärosätens antagning till högskoleutbildning.⁵⁶ Vid den starkt decentraliserade antagningen till högskolan spelar även annat än betygen roll. Sökanden behöver vanligen bifoga personligt brev och rekommendationsbrev, men urvalsprocessen kan även omfatta intervjuer, muntliga och skriftliga tester, presentationer och bedömningar av GCSE-resultat. För sökande direkt från gymnasieskolan används lärarbedömningar av förväntade *A-levels*-resultat för de elever som ännu inte har skrivit proven. Antagningen kan dock villkoras med att eleven uppnår den förväntade nivån.

För den som efter grundskolan inte vill studera vidare med sikte på högre utbildning finns även möjlighet till yrkesutbildning eller lärlingsutbildning efter GCSE. Sedan år 2020 finns även *T-levels* som en möjlighet till yrkesutbildning med arbetslivsanknytning och om-

⁵³ Se t.ex. Cambridge University Press & Assessment (2022, 2024) och Think Student (2022).

⁵⁴ Meadows (2016).

⁵⁵ Ofqual (2017).

⁵⁶ UCAS (2024).

fattande praktik samtidigt som den är förberedande för viss högre utbildning.⁵⁷

Estland⁵⁸

Skol- och betygssystem

Bedömning och betygssättning regleras i den estniska läroplanen, där bedömningen beskrivs som en integrerad del i undervisningen och utbildningen i den nioåriga grundskolan.⁵⁹

Lärarsatta betyg ges enligt lag på en femgradig skala där 3 vanligen betecknas som godkänt (5 mycket god, 4 god, 3 tillfredsställande, 2 inte så god och 1 svag).⁶⁰ Skolor får använda andra skalor eller verbala omdömen till och med årskurs 5. Det betygssystem som används och principerna för omräkning av betygen till den femgradiga skalan ska anges i skolans lokalt framtagna läroplan. När eleven lämnar skolan ska de sammanfattande betygen för det läsåret och de betyg som erhållits under det pågående kvartalet omvandlas till en femgradig skala. Om bara muntliga omdömen har använts har eleven rätt till ett numeriskt betyg i slutet av årskurs 6 eller om eleven byter skola. Varje skola ska publicera en egen plan för undervisning, bedömning och betygssättning. Läroplanen anger att skriftliga uppgifter ska bedömas så att 90–100 % rätt ger 5 i betyg, 75–89 % rätt ger betyg 4, 50–74 % rätt ger betyg 3, 20–49 % rätt ger betyg 2 och 0–19 % rätt ger betyg 1.⁶¹

I början av årskurs 4 och 7 gör tio procent av årskohorten digitala prov i estniska och matematik i syfte att mäta kunskaper nationellt. Skolor som inte valts ut för dessa prov kan välja att göra dem ändå för att följa upp sina egna elever. Elever får inte provbetyg utan ett skriftligt omdöme.

⁵⁷ Department for Education (2023).

⁵⁸ Huvudsaklig källa för avsnittet: e-postkontakt med Estlands Education and Youth Board och myndighetens hemsida, <https://harno.ee/en>.

⁵⁹ Riigi Teataja (2011).

⁶⁰ Riigi Teataja (2010).

⁶¹ Riigi Teataja (2011).

High stakes-prov i grundskolan och gymnasieskolan

I slutet av årskurs 9 gör eleverna *basic school final examinations* som innefattar ett specialarbete samt tre slutprov i estniska, matematik och ett valfritt ämne som de måste klara för att få ett avgångsbetyg och bli behöriga till gymnasiestudier.⁶² Enligt den nationella läroplanen ska eleverna innan de skriver examensproven ha fått sina lärarsatta betyg. Alla prov är skriftliga utom prov i främmande språk som även innehåller ett muntligt delprov. De valfria ämnen som elever kan välja mellan är ryska, biologi, kemi, fysik, geografi, historia, samhällskunskap, engelska, franska, tyska eller ryska som främmande språk.⁶³

Staten eftersträvar att nationella prov ska vara digitala och kunna rättas automatiskt i så hög grad som möjligt, men mycket bedöms fortfarande av lärare på skolor. Lärarna som bedömer får anonymiserade provfrågor sorterade uppgiftsvis, dvs. de ser varken elevens namn eller någon enskild elevs prov i sin helhet. Riktlinjer för bedömning finns för varje prov, men uppföljning visar att bedömaröverensstämmelsen är låg. Central manuell bedömning förekommer för vissa prov i exempelvis språk och då är det ofta två personer som bedömer varje prov. Bedömningen kan överklagas.

Estniska gymnasieskolor får själva bestämma på vilka grunder de antar elever. Således är inte godkänt i alla ämnen i grundskolan ett formellt krav för antagning. Vanligast är att skolor arrangerar egna antagningsprov och dessa hålls ofta på våren innan examensproven har ägt rum. Elever skriver ofta flera prov eftersom de ansöker till flera gymnasieskolor. Det förekommer också att slutbetyg från grundskolan eller resultaten på de nationella proven i årskurs 9 används som urvalsgrund.

Den högskoleförberedande gymnasieskolan är kursbaserad och ska omfatta minst 96 kurser à 35 undervisningstimmar (45 minuter).⁶⁴ Gymnasieskolan ska använda samma femgradiga betygsskala som grundskolan men under utbildningens gång används ofta en 100-gradig skala vid examinationer av kurser. Först i slutexamen måste skolan konvertera poängen och använda den femgradiga betygsskalan. Ungefär en fjärdedel av de gymnasiestuderande väljer yrkesförberedande

⁶² Haridus- ja Noorteamet (2021).

⁶³ Haridus- ja Noorteamet (2021).

⁶⁴ Haridus- ja Teadusministeerium (2024).

utbildningar och där är examensproven frivilliga.⁶⁵ Där används en fyrgradig skala med betyg från 2 till 5 där 2 är underkänt. I flera yrkesämnen finns dessutom bara betygen godkänt och underkänt.

Högskoleförberedande gymnasieutbildningar avslutas med examensprov i estniska, engelska och matematik. Examensproven ges av staten vid två tillfällen på våren och elevens provresultat, tillsammans med de lärarsatta betygen, spelar vanligtvis roll vid antagning till högre utbildning. Båda publiceras parallellt i examensbeviset. Lärosäten arrangerar ofta egna antagningsprov eller intervjuer som kompletterande urvalsgrund för antagning till högskolan.

Både *basic school final examinations* och examensproven i årskurs 9 konstrueras av den estniska utbildningsmyndigheten, Education and Youth Board. Bedömning av examensprov kan överklagas till utbildningsdepartementet där en panel granskar att bedömningen har gjorts i enlighet med bedömningsmanualen.⁶⁶ Ombedömningen kan resultera i oförändrat, höjt eller sänkt provresultat. Det förekommer i dagsläget inte ankaruppgifter i eller åtgärder för ekvivalering av proven men viss utprövning av provuppgifter görs. Ett arbete för att utveckla provkonstruktionen och bedömningen har dock påbörjats.

Frankrike⁶⁷

Grundskola och betygssystem

Grundskolan består av *école maternelle* (3–5 år), *école élémentaire* (6–10 år) och *collège* (11–14 år) och är obligatorisk. Elevens prestationer under grundskolan återkopplas regelbundet till vårdnadshavarna, i en särskild bok, *livret scolaire*. Det finns nationellt framtagna formulär för detta. Lärarbetygssättningen är inte formellt reglerad vad gäller betygsskala eller årskurser för första betyg. Lärare uppmanas att sätta betyg så sent som möjligt i de första cyklerna.

Nationella diagnostiska prov görs i slutet på grundskolans två första cykler. Dessa prov har ingen betygsstödjande funktion och ska inte beaktas av lärare för annat än för att ge stödinsatser.

⁶⁵ Haridus- ja Noorteamet (2024).

⁶⁶ Haridus- ja Noorteamet (2024).

⁶⁷ Huvudsakliga källor för avsnittet: e-postkontakt med Ministère de l'Éducation Nationale et de la Jeunesse, avdelning DREIC B-2 och det franska utbildningsdepartementets hemsida, <https://www.education.gouv.fr/>.

Betyg ges från och med *collège* på en 1–20-skala men vanligen används endast stegen 10–17. Även om det inte är formellt reglerat så betecknas 10 generellt som godkäntgräns. De allra högsta stegen används endast för exceptionella prestationer. Det finns inga krav på specifika betyg i vissa ämnen eller genomsnittsbetyg för att få börja nästa årskurs utan bedömningar av eventuella behov av att gå om en årskurs görs av lärare på klasskonferenser.

Grundskoleexamen, *brevet*

Grundskolan avslutas med examen, *brevet*, för elever som får minst 400 av 800 i totalbetyg på följande:

- skriftliga prov i franska (100 p), matematik (100 p), ett SO-ämne (50 p) och ett NO-ämne (50 p),
- en muntlig presentation (100 p) samt
- lärarbedömningar av elevens kompetens på åtta så kallade kärnområden⁶⁸ (400 p varav bedömningen sker på en fyrgradig skala: 10 otillräckliga, 25 svaga, 40 tillräckliga, 50 mycket goda kunskaper).⁶⁹

Utöver godkäntgränsen, 400 p, kan eleven få särskilda utmärkelser om de uppnår 480, 560 respektive 640 p och de kan få bonuspoäng för att ha läst extra valfria ämnen såsom språk.

Brevet är från och med år 2025 ett krav för tillträde till gymnasieskolan, *lycée*, och yrkesutbildningar. Antagningen till den offentligfinansierade gymnasieskolan görs framför allt på grundval av geografisk närhet.⁷⁰ Hänsyn tas även till särskilda behov, social sammansättning och utbildningsbakgrund och det förekommer att rekommendationsbrev beaktas. När det är fler sökande än platser tillämpas vanligen ett system för att rangordna eleverna baserat på resultaten i det åtta kärnområdena och andra studieresultat. Privata gymnasieskolor kan ha egna urvalsgrunder.

⁶⁸ Humanistiska värden, språklig kompetens, grundläggande matematisk och vetenskaplig kompetens, digital kompetens, personligt projekt samt rättigheter och skyldigheter som medborgare, skriftlig och muntlig kommunikation, globala frågor, fysisk aktivitet och hälsa.

⁶⁹ Ministère de l'Éducation Nationale et de la Jeunesse (2024a).

⁷⁰ Ministère de l'Éducation Nationale et de la Jeunesse (2024b).

Gymnasieskola och studentexamen, baccalauréat

Den franska gymnasieskolan är treårig, består av årskurserna *la seconde*, *la première* och *la terminale* och ger en studentexamen: *baccalauréat*, ibland förkortat *bac*. Studentexamen instiftades av Napoleon år 1808. Det finns tre olika examina: en generell med olika högskoleförberedande inriktningar, en som kallas teknisk som förbereder för kortare yrkesinriktad påbyggnadsutbildning och en variant som ger yrkesexamen på gymnasial nivå.

För att få en studentexamen genomför eleverna skriftliga och muntliga prov.⁷¹ Även för *baccalauréat*-examen är betygsskalan 1–20: 16–20: mycket bra (*très bien*: TB); 14–15: bra (*bien*: B); 12–13: tillfredsställande (*assez bien*: AB); 10–11: godkänd (*passable*; inget officiellt betyg dock); 0–9: underkänd (*insuffisant*). *Bac* är normalt sett behörighetsgivande för högre studier men antagning och principer för urval till olika utbildningar varierar. Andelen elever som fått de högsta betygen har ökat de senaste åren.⁷²

Proven för baccalauréatexamen

År 2021 reformerades de franska examensproven så att resultaten på de olika slutproven ges olika vikt. Tillsammans står de numera för 60 % av elevens slutbetyg. Resterande 40 % av slutbetyget består i stället av lärarens kontinuerliga bedömning av elevens prestationer under de två sista gymnasieåren (*première* och *terminale*).⁷³

Examensproven genomförs i de två sistnämnda årskurserna i ämnena franska, filosofi och två valfria ämnen. Dessutom görs en muntlig presentation där eleven presenterar ett genomfört undersökande arbete för två examinatorer. Dessa ska inte vara tidigare kända för eleven. Slutproven i franska görs i slutet av *première*, dvs. det näst sista läsåret.⁷⁴

Det förekommer att vissa provdelar skannas in för att automaträttas. Det går inte att överklaga bedömningen bortsett från om man upptäcker felaktigheter i rättningen. Skriftliga examensprov bedöms dock i huvudsak anonymiserade och av lärare som inte har

⁷¹ Ministère de l'Education Nationale et de la Jeunesse (2024a).

⁷² Beau (2022).

⁷³ Ministère de l'Education Nationale et de la Jeunesse (2024c) och Ministère de l'Education Nationale et de la Jeunesse (2021).

⁷⁴ Ministère de l'Education Nationale et de la Jeunesse (2024d).

undervisat den aktuella klassen. Det är obligatoriskt för lärare att delta i arbetet med proven. Huvuddelen av frågorna är av essä- eller problemlösningsskaraktär och bedöms manuellt. Som stöd för lärarna finns skriftliga bedömningsstöd med principer och rekommendationer i syfte att uppnå allsidig och likvärdig bedömning. Det finns ingen årsvis standardisering eller princip om viss fördelning av betyg.

De årliga medelbetygen i elevens olika kurser används för en del av slutbetyget, men precis som för proven viktas olika ämnen olika tungt.

Eftergymnasial utbildning

Det franska systemet för högre utbildning kännetecknas av en bredd av offentliga och privata lärosäten, som regleras av olika regelverk och lagstiftning, inom följande huvudkategorier:

- offentliga universitet, *universités*, och *COMUEs* (*Communautés d'Universités et Établissements*),
- övriga offentliga och privata högskolor och institut utanför universiteten samt
- offentliga *grandes écoles*.

Samtliga kräver normalt gymnasieexamen, *baccalauréat*, för behörighet, men det i sig är sällan tillräckligt för tillträde. På offentliga franska universitet kan den som är behörig ofta registrera sig på enskilda kurser eller längre utbildningar på kandidatnivå utan att behöva söka i konkurrens. Utbildningarna är i stort sett offentligfinansierade, men lärosätena kan ta ut mindre s.k. registreringsavgifter av studenterna. Traditionen lever kvar att alla som uppfyller kraven och önskar studera erbjuds en utbildningsplats.⁷⁵ En stor ökning av sökande över tid har dock omöjliggjort att alla kan få en plats på sitt förstahandsval⁷⁶ och under en period förekom fördelning av platser enligt principen först-till-kvarn och genom lottning.⁷⁷ Ansökan till grundutbildning görs numera centraliserat till många utbildningar och lärosäten, men de sökande kan inte rangordna sina val. Lärosätena specificerar vad ansökan ska innehålla och

⁷⁵ Lagen om högre utbildning, *Article 14 de la loi du 26 janvier 1984 sur l'enseignement supérieur*.

⁷⁶ Graveleau (2017).

⁷⁷ Parcoursup (2024).

sköter själva urval och antagning.⁷⁸ Systemet, *Parcoursup*, har fått kritik för både sin bristande tekniska funktionalitet och att inte vara utformat användarvänligt eller transparent för de sökande.⁷⁹ Förstågångssökande, som söker medan de fortfarande går i gymnasieskolan, söker och bedöms baserat på de examensprov som de skrivit och sina lärarsatta gymnasiebetyg eftersom de ännu inte har slutfört sin *bac*.

Det första årets universitetsstudier består ofta i huvudsak av storföreläsningar och självstudier inför omfattande tentamina som avgör vilka, betydligt färre, som får fortsätta till nästa, mer specialiserade, nivå. Systemet syftar till jämlik tillgång till utbildning men har kritiserats för att vara ett kostsamt, och samtidigt underfinansierat, ett år långt antagningsprov där bara hälften klarar sig vidare. Därför tillämpar en del lärosäten numera viss begränsning och lottning.⁸⁰ Även masterprogram söks direkt hos lärosätena som själva avgör vilka urvalsprinciper för antagning som ska tillämpas.

Det finns tvåårig yrkesinriktad eftergymnasial utbildning, *brevet de technicien supérieur (BTS)*, som ofta ges av eller i samarbete med privata eller offentliga gymnasieskolor.⁸¹ Andra former av privat avgiftsbelagd högre utbildning finns på yrkesinriktade områden såsom teknik och handel. I övrigt-kategorin finns också katolska lärosäten. Dessa får inte utfärda kandidat-, magister- eller doktors-examina utan behöver samarbeta med andra universitet för det. Katolska universitet kan dock i eget namn utfärda diplom inom alla ämnesområden, samt kanoniska diplom inom kyrkovetenskap, kanonisk rätt, teologi och filosofi.⁸² Till övriga räknas även privata skolor på gymnasial nivå som erbjuder högskoleförberedande kurser till de sistnämnda i listan, *les grandes écoles*.

⁷⁸ Parcoursup (2024).

⁷⁹ Stromboni (2018).

⁸⁰ Lichfield (2015).

⁸¹ Ministère de l'Enseignement Supérieur et de la Recherche (2020).

⁸² Décret n° 2009-427, 16 april 2009.

High stakes-prov för antagning till *grandes écoles*: concours

Grandes écoles är en specifik fransk utbildningsform som främst antar studenter baserat på resultaten på olika skriftliga och muntliga prov, intervjuer, rekommendationsbrev och arbetsprover, s.k. *concours*.⁸³ Antagningsprocessen är decentraliserad till nämnder på respektive lärosäte men det finns ett nationellt forum för rekryteringsfrågor.⁸⁴ Till skillnad från till *universités* tillämpas en höggradigt selektiv och icke-transparent antagning. De som antas, cirka 5 % av landets studentpopulation, säkerställer, redan genom att de passerar nålsögat, goda karriärmöjligheter i staten eller privat sektor, i Frankrike eller utomlands.⁸⁵ Även om det inte är ett krav, har de som antas nästan alltid två eller tre år av krävande, förberedande kurser (*classes préparatoires* eller *prépa*) i bagaget.⁸⁶ Dessa kurser ses som eftergymnasiala men brukar ges på privata och offentliga *lycées*, dvs. gymnasieskolor. Utbildningsanordnarna får själva besluta om antagnings- och urvalsprinciper, men resultaten från *bac* beaktas oftast i någon utsträckning. Syftet med *prépa* är uttalat att *teach-to-the-test* så att studenterna passerar *concours* till *les grandes écoles*.⁸⁷ Många antas ändå inte och trots kurserna, som gett dem språkkunskaper, omfattande studievana, analysförmåga och testat förmågan att organisera arbete under press, får de varken examen eller jobb. Därför ger de numera ECTS-poäng och en del *universités* har beslutat att *prépa*-kurser kan tillgodoräknas som del av en reguljär universitetsutbildning.

Grandes écoles är i allmänhet offentligt finansierade – vissa betalar till och med studielön – samtidigt som de har nära samarbeten med näringslivet – och de som kallas *grands établissements* har i regel både utbildning och forskning.⁸⁸ Vissa specialiserade *grandes écoles*, t.ex. inom konst, arkeologi och arkitektur, kräver inte förberedande utbildning utan har egna antagningsprov direkt efter *bac*.

⁸³ Inte att förväxla med *concours général* eller särskilda *concours* för befordran av lärare och disputerade. *Concours général* är särskilda skololympiader i 30 ämnen för elever på gymnasienivå. Elever med särskilt goda resultat i dessa kan dock i vissa fall antas till *grandes écoles* direkt från gymnasiet. Se t.ex. <https://www.education.gouv.fr/le-concours-general-des-lycees-et-des-metiers-un-prix-d-excellence-10022> och Olympic Scholarships | ENS (psl.eu).[2024-07-16].

⁸⁴ Conférence des Grandes Écoles (2023).

⁸⁵ Lichfield (2015).

⁸⁶ Ministère de l'Éducation Nationale et de la Jeunesse (2013).

⁸⁷ Calmand m.fl. (2009).

⁸⁸ Eurydice (2023a).

Nederländerna⁸⁹

Skol- och betygssystem

I Nederländerna omfattar den obligatoriska skolan, *Basisonderwijs*, minst ett år på förskola och därefter ytterligare sex år i grundskolan. Vid 12 års ålder börjar eleverna vid den lägre sekundärskolan och redan då väljer de mellan en yrkesförberedande fyraårig utbildning (VMBO), en femårig allmän utbildning (HAVO) eller en sexårig universitetsförberedande utbildning (VWO). De flesta sekundärskolorna erbjuder flera inriktningar. En utgångspunkt är att det ska vara enkelt för eleverna att byta inriktning. Det går även att börja *praktijkonderwijs*, en praktisk femårig utbildning av lärlingskaraktär, för den som bedöms ha svårigheter att klara av annan utbildning.

Skolan avgör själv om en elev bedöms kunna flyttas upp till nästa årskurs.⁹⁰ Kritik mot den tidiga uppdelningen och skolornas stora inflytande över elevernas utbildning finns och en studie vid universitetet i Groningen har exempelvis visat ett starkt samband mellan lägre föräldrainsomst och råd till elever att välja lägre utbildning än vad deras testresultat indikerar som lämpligt.⁹¹

Betyg sätts vanligtvis inte i den första delen av den obligatoriska grundskolan. De flesta grundskolor beskriver varje elevs framsteg inom olika ämnesområden i en rapport som sammanställs en gång per termin. Vissa skolor rapporterar om elevernas framsteg med hjälp av poäng, medan andra använder beskrivande indikatorer. Skolorna kan själva bestämma exakt hur de mäter elevernas utveckling. Sedan läsåret 2014/2015 är det obligatoriskt för skolor att genomföra ett kunskapstest för avgångselever från grundskolan och använda ett elevuppföljningssystem (LVS).⁹²

⁸⁹ Huvudsakliga källor för avsnittet: e-postkontakt med *Examenloket*, College voor Toetsen en Examens (Exam board of the Ministry of Education in the Netherlands) samt *Examenbladmbno.nl* (Board of Tests and Examinations).

⁹⁰ Eurydice (2024d).

⁹¹ KansenKaart, indikator *Schooladvies lager dan eindtoetsadvies*, <https://kansenkaart.nl/schooladvieslager>.

⁹² Eurydice (2024d).

Nationella prov på grundskolenivå

En särskild del av det nederländska skolsystemet är att eleverna regelbundet gör diagnostiska tester i språk, matematik och studieteknik. Testerna har sedan 1968 utvecklats av ett nationellt test- och bedömningsföretag, Cito. Testresultaten visar enskilda elevers resultat i jämförelse med jämnåriga men används även för jämförelser mellan skolor i syfte att dessa ska kunna justera sin undervisning utifrån vad testresultaten indikerar.⁹³

I den lägre sekundärskolan görs nationella slutprov och ibland har skolorna även egna examensprov i vissa ämnen. I praktisk-estetiska ämnen ges godkänt/icke godkänt i stället för sifferbetyg och i SO-ämnen ges inte nationella prov. Flera provtillfällen brukar erbjudas under året. Resultaten på de olika proven (50 %) viktas med lärarsatta betyg (50 %) och utgör tillsammans elevens slutbetyg. Slutbetyget används inte för ansökan till den högre sekundärskolan utan det fortsatta utbildningsvalet avgörs av lärares rekommendation om fortsatt studieinriktning.

Utbildning på gymnasial nivå och examensprov, *Eindexamen*

Den högre sekundärskolan påbörjas vid cirka 15 års ålder och eleverna inriktar sig då allt tydligare på sin sedan tidigare valda studieinriktning om de inte rekommenderats att byta inriktning. I slutet av de högskoleförberedande utbildningarna genomförs *Eindexamen*, som består både av nationella prov och skolornas egna examensprov på de högskoleförberedande inriktningarna. Departementet för utbildning, kultur och vetenskap föreskriver vilka ämnen som måste undervisas under examensåret.⁹⁴ Det ges nationella skriftliga examensprov i flera ämnen för alla elever som går samma typ av utbildning och det förekommer både digitala prov och prov på papper. De nationella proven konstrueras av departementet för utbildning, kultur och vetenskap och äger rum i slutet av maj, med möjlighet till omprov i juni och augusti.⁹⁵ Hjälpmedel som miniräknare och vissa uppslagsböcker är tillåtna.

⁹³ Cito har avböjt att besvara detaljerade frågor från utredningen.

⁹⁴ Government of the Netherlands (u.å.).

⁹⁵ Eurydice (2023b).

Skolorna bestämmer därutöver själva vilka ämnen som eleverna ska skriva prov i. Varje år måste skolorna lämna in sin provplan till National Institute for Curriculum Development (SLO) som i sin tur publicerar riktlinjer för prov för varje ämne och utbildningsnivå. Proven består vanligtvis av två eller flera muntliga, praktiska eller skriftliga delprov per ämne och ska ha slutförts innan de nationella proven ges. Ämnen som inte ingår i den nationella examensramen kan även slutföras före det sista skolåret. Elever kan göra prov i extra ämnen om de så önskar och betyg i extra ämnen beräknas på samma sätt som för övriga provämnerna. En elev kan också välja att göra prov på högre nivå än vad de har studerat för. En HAVO-elev kan t.ex. göra engelskprovet på den högskoleförberedande VWO-nivån. För elever med funktionshinder, dyslexi, kronisk sjukdom, eller om ens förstaspråk inte är nederländska, kan extra anpassningar såsom tidstillägg eller ytterligare medhavd ordbok medges.⁹⁶ Den som på grund av sjukdom eller av andra oförutsedda skäl inte kunnat skriva examensproven får göra det inom två år.⁹⁷ De nationellt centralt framtagna proven rättas huvudsakligen automatiskt. För de prov som rättas manuellt bedöms de först av elevernas egna lärare och skickas sen för en andra bedömning av andra lärare. Rättningen kan inte överklagas.

Betygsskalan för de olika proven såväl som slutbetyget är 1–10. Gränsen för godkänt på de studieförberedande utbildningarna är 5,5, men det finns viss möjlighet till att ha ett eller två ämnesbetyg som är 4 eller 5 i vissa ämnen.⁹⁸ Slutbetyget i varje ämne är genomsnittet av betyget på skolans examensprov och betyget på det nationella skriftliga provet och det avrundas till heltal. För ämnen där det endast finns ett skolprov är det erhållna betyget slutbetyget. I praktisk-estetiska ämnen ges godkänt/icke godkänt i stället för sifferbetyg.⁹⁹

Antagning till högskolan

För antagning till högskola (HBO) kan utbildning från HAVO, VWO eller VMBO ge behörighet. Till forskningsinriktade universitet (WO) krävs däremot en VWO-examen eller ett års förberedande

⁹⁶ Government of the Netherlands (u.å.).

⁹⁷ Government of the Netherlands (2023).

⁹⁸ Eurydice (2023b).

⁹⁹ Förordning om examination, Uitvoeringsbesluit WVO 2020, Staatsexamen.

studier på ett HBO-program. Det finns ett gemensamt ansöknings-system för lärosäten, men såväl antagningen som tillämpningen av urvalsgrunder är decentraliserad. Till vissa utbildningar antas alla sökande som uppfyller behörighetskraven. Exempelvis får studenter med ett genomsnittsbetyg på åtta eller högre i sin studentexamen automatiskt en plats på en sökt HBO-utbildning.¹⁰⁰ Det förekommer dock att lärosäten håller *study choice tests* i början av första terminen, i syfte att direkt sälla ut studenter som inte förväntas kunna fullfölja utbildningen. För antagning till vissa populära utbildningar, t.ex. inom det medicinska området, tillämpas en begränsning av antalet studerande (*numerus fixus*) och platserna fördelas baserat på en rangordning av de sökande. De sökande får göra tester, skriva hemuppgifter och personliga brev som varierar beroende på sökt utbildning och lärosäte. Det förekommer att betyg från gymnasial utbildning är del av de urvalsgrunder som viktas för rangordningen.¹⁰¹

Polen¹⁰²

Skol- och betygssystem

Grundskolan är åttaårig och påbörjas vid sju års ålder efter ett års obligatorisk förskola. Redan i förskolan ges eleverna ett skriftligt omdöme inför skolstarten. Läraromdömen sätts även varje termin i de lägre årskurserna. Från och med årskurs 4 ges sifferbetyg enligt skalan i tabell 1:

Tabell 1 Polens betygsskala

Sifferbetyg	Beskrivning	Omräkning till antagningspoäng vid examen
1	Underkänt	0
2	Acceptabelt	2
3	Tillräckligt	8
4	Bra	14
5	Mycket bra	17
6	Utmärkt	18

Källa: Polens förordning om utbildning, 18 november 2022.

¹⁰⁰ Eurydice (2023c).

¹⁰¹ Exempel på universitet som tillämpar rankning där gymnasiebetyg kan utgöra en del i urvalsgrunden: Radboud University, Erasmus University, Rotterdam och University of Amsterdam.

¹⁰² Huvudsakliga källor för avsnittet: muntlig information från Central Examinaton Board (CKE).

En elev som har underkänt i fler än tre ämnen behöver, i augusti inför kommande läsår, i skriftliga prov visa att den har tillräckliga kunskaper för att börja nästa årskurs, annars får den gå om en årskurs.

High stakes-prov i grundskolan

Årskurs 8 avslutas med obligatoriska skriftliga examensprov, *egzamin ósmoklasisty*, i polska, matematik och främmande språk. Inga prov är digitala. Resultaten på de tre examensproven står tillsammans för hälften av elevens slutbetyg (polska och matematik 35 % vardera och främmande språk 30 %).¹⁰³ Resten utgörs av de lärarsatta ämnesbetygen men elever kan också få extra poäng för deltagande och resultat i kunskapsstävlingar, *olympiader*, eller aktiviteter utanför det obligatoriska schemat. De sammantagna slutbetygen omräknas till maximalt 200 antagningspoäng, vilket är den huvudsakliga grunden för urval till gymnasieskolan. Examensproven ges i maj och juni. För den som har giltig frånvaro från ett examensprov dubblas i stället det lärarsatta betyget.

Gymnasieskola och examensprov

Det krävs inte godkända resultat på examensproven eller i alla ämnen för att antas till fortsatt skolgång men grundskolan behöver intyga att eleven har genomgått utbildningen och olika gymnasieskolor hanterar antagningen olika. Elever med särskilt goda resultat i olympiader behöver t.ex. inte göra examensproven utan får platser på sina önskade gymnasieutbildningar. Det finns generell, *liceum*, (fyra år) eller kombinerad yrkesinriktad/högskoleförberedande, *technikum*, (fem år) gymnasieskola.¹⁰⁴ På gymnasial nivå finns också två- respektive treårig yrkesskola som kan avslutas med särskilda yrkesprov. Efter grundskolan är det utbildningsplikt till 18 års ålder men den som inte vill gå i gymnasieskola eller yrkesutbildning kan få arbetsplatsförlagd lärlingsutbildning. Den som gått en högskoleförberedande utbildning men ändrat sig kan komplettera med ett eller två år på yrkesskola för att få en yrkesexamen. Ungefär hälften av eleverna på gymnasial nivå i Polen går en yrkesinriktad utbildning.

¹⁰³ Polens förordning om utbildning, 18 november 2022.

¹⁰⁴ UNESCO-UNEVOC (2022).

I gymnasieskolan ges lärarsatta betyg, skala 1–6, varje termin precis som i grundskolan. Det går att avsluta gymnasieskolan med enbart ett samlat betygsdokument men den som vill söka till högskolan behöver göra skriftliga och muntliga examensprov, *egzamin maturalny*, i polska, matematik och moderna språk samt i ett till sex valfria ämnen. Även på yrkesskolorna, som inte är högskoleförberedande, görs skriftliga och praktiska examensprov. Vanligen krävs godkända prov i ett till tre av 270 tillgängliga yrkesämnen. Det gäller att få 50 % rätt på de skriftliga och 75 % rätt på de praktiska proven för att bli godkänd. Enskilda skolors provresultat är offentliga och utbildningsdepartementet använder resultaten från proven i årskurs 8 och på maturaproven som en indikator på skolors kvalitet.

Antagning till högskoleutbildning ska enligt lag baseras till minst 50 procent på de sökandes maturaproverresultat. Lärosätena får använda andra kompletterande urvalsgrunder också men de får inte ha antagningsprov som prövar samma saker som maturaproven.

Maturaproven¹⁰⁵

Examensprov erbjuds för närvarande i 26 ämnen. Det finns muntliga provdelar i språk men i övrigt är det skriftliga prov och inga provdelar är digitala. Valet av vilka prov en elev skriver bestäms oftast av vilken utbildning hon eller han avser att söka vidare till. Lärosäten ska i god tid, cirka 1,5 år i förväg, publicera vilka behörighetskrav som krävs för respektive utbildning så att elever kan planera sina examensprov utifrån detta.

Proven konstrueras centralt av myndigheten *Centralna Komisja Egzaminacyjna*, CKE. De innehåller ankaruppgifter och utprövas i syfte att standardisera svårighetsgraden över tid. Flera prov kan skrivas på grundläggande eller avancerad nivå. Alla prov ska finnas anpassade för elever med funktionsnedsättningar. I dagsläget finns generella riktlinjer för examensprov men från och med år 2025 ska proven konstrueras så att de går i linje med respektive ämnesplan.¹⁰⁶ CKE ansvarar också för att publicera bedömningsanvisningar. Bedömningen organiseras regionalt och sker således anonymt och externt. Verksamma lärare som vill arbeta som bedömare behöver gå

¹⁰⁵ Centralna Komisja Egzaminacyjna (2024a).

¹⁰⁶ Centralna Komisja Egzaminacyjna (2024b).

en årlig utbildning. Sedan 2015 kan elever överklaga bedömningen, först till regional och sen till nationell nivå. Överklagandemöjligheten uppges ha gjort det svårare att rekrytera bedömare. För att bli godkänd krävs 30 procent rätta svar i de tre obligatoriska ämnena. Det finns ingen godkäntgräns för de valbara ämnena.

Nordrhein-Westfalen (Tyskland)¹⁰⁷

Skol- och betygssystem

Nationellt framtagna kunskapsstandards, *Bildungsstandards*, specificerar vad tyska delstater (förbundsländer) ska ha för läroplans-element för kärnämnen och fungerar som bindande mål för alla delstater. Läroplanerna bestäms och hanteras i övrigt av respektive delstat och som ett resultat av decentraliseringen har nästan varje delstat egna skolformer och läroplaner för ämnen och årskurser.¹⁰⁸

I Tysklands folkrikaste delstat, Nordrhein-Westfalen, är grundskolan, *grundschule*, fyraårig och påbörjas vid sex års ålder.¹⁰⁹ Skolplikten varar till 18 års ålder eller slutförd gymnasieutbildning.¹¹⁰ Från och med femte årskursen påbörjas sekundärskolan och det finns flera olika fem- respektive sexåriga skolformer som del av *sekundarstufe I* att välja mellan:

- *Hauptschule*, sex år förberedande för yrkesutbildning,
- *Realschule*, sexårig, i huvudsak förberedande för utbildning med yrkesinriktning men med teoretiska ämnen som även möjliggör högskoleförberedande gymnasieutbildning,
- *Gesamtschule*, sexårig både teoretisk och praktisk utbildning med undervisning på två nivåer (grundläggande respektive utvidgad) i flera ämnen, i en skolform som även omfattar fortsatt utbildning på gymnasial nivå,
- *Sekundarschule*, sexårig både teoretisk och praktisk utbildning men fortsätter inte till gymnasial nivå, och

¹⁰⁷ Huvudsakliga källor för avsnittet: muntlig information från IQB.

¹⁰⁸ Wendt m.fl. (2016).

¹⁰⁹ Ministerium für Schule und Weiterbildung des Landes Nordrhein-Westfalen (2015).

¹¹⁰ Ministerium für Schule und Weiterbildung des Landes Nordrhein-Westfalen (2015).

- *Gymnasium*, femårig studieförberedande utbildning (ej att förväxla med utbildning på gymnasial nivå).

Alla elever flyttas upp från årskurs 1 till årskurs 2. Därefter förekommer det att elever får gå om. Inga nationella prov, betyg eller liknande används dock för formell uppflyttning eller ansökan till högre årskurs. Lärare bedömer elevernas kunskaper utifrån en kontinuerlig kunskapsutvärdering och sätter vanligen betyg från årskurs 3, på skala 1–6 där 1 är bäst och 5 och 6 är underkända betyg.¹¹¹ I många tyska delstater är det upp till läraren att förorda huruvida elever bör gå om, i synnerhet om de har getts betyget 6 i något ämne. Läraren rekommenderar också val av skolform från och med årskurs 5. Betygen är målrelaterade.

Tabell 2 Exempel på tysk betygsskala som används på vissa utbildningar på gymnasial nivå

Vitsord	Betyg	Poäng	Beskrivning
Mycket bra	1	15–13	Prestationerna uppfyller kraven i särskilt hög grad.
Bra	2	12–10	Prestationerna uppfyller kraven fullt ut.
Tillfredsställande	3	9–7	Prestationerna uppfyller kraven generellt.
Tillräckligt	4	6–5	Prestationerna har vissa brister men uppfyller kraven i stort.
Knappt tillräckligt	5	4	Prestationerna är bristfälliga och uppfyller endast i begränsad utsträckning kraven.
Otillräckligt	6	3–1	Prestationerna uppfyller inte kraven, men visar att nödvändiga grundläggande kunskaper finns och att bristerna kan avhjälpas inom överskådlig framtid.
Otillfredsställande	–	0	Prestationerna uppfyller inte kraven; även de grundläggande kunskaperna är så bristfälliga att bristerna inte kan avhjälpas inom överskådlig framtid.

Källa: Förordning, *Verordnung über den Bildungsgang und die Abiturprüfung in der gymnasialen Oberstufe (APO-GOST)*, från 1998, ändrad 2023.

VERA-proven

Alla tyska delstater, utom Niedersachsen där det är frivilligt för lärare att bestämma om deltagande, har sedan läsåret 2007/2008 skriftliga nationella prov, *Vergleichsarbeiten*, VERA, i matematik och tyska

¹¹¹ Skollagen för Nordrhein-Westfalen, Schulgesetz für das Land Nordrhein-Westfalen (Schulgesetz NRW - SchulG) från 2005, ändrad 2022.

i årskurs 3 samt i matematik, tyska och främmande språk i årskurs 8.¹¹² Proven konstrueras av Institut zur Qualitätsentwicklung im Bildungswesen, IQB, ett institut knutet till Humboldtuniversitetet, och ska vara standardiserade diagnostiska prov baserade på *Bildungsstandards*. I Nordrhein-Westfalen kommuniceras proven dock som ”*learning status surveys*”.¹¹³ Proven används inte som grund för betygssättning utan resultaten ska ge lärarna information om elevernas styrkor och svagheter samt ämnesspecifika pedagogiska och pedagogiska psykologiska rekommendationer för att hjälpa dem att förbättra sin undervisning.

IQB tar fram provfrågor som sen prövas ut och utvärderas av lärare över landet. Proven bedöms av lärare på den egna skolan. I vissa delstater är proven dock digitala och där kan delar av dem automaträttas.¹¹⁴ Skolorna förses med nivåbeskrivningar för bedömning av enskilda elevers provresultat. En beskrivning omfattar vanligtvis fem kompetensnivåer (I–V), där nivå III representerar den nivå som eleverna i genomsnitt förväntas nå.¹¹⁵ IQB publicerar inte skolors VERA-resultat för jämförelse utan menar att skolor själva ska arbeta med sina resultat på klassnivå. Med ungefär fem års mellanrum gör IQB fördjupade jämförande analyser av VERA-resultaten på nationell nivå. Därtill gör IQB trendmätningar, *IQB-Bildungstrend*.

Gymnasieskola och examensprov, *Abitur*

Utbildningen på gymnasial nivå som följer är fyraårig, *Sekundarstufe II*, och det finns både *Gymnasiale Oberstufe* som är högskoleförberedande och *Berufskolleg* för yrkesutbildning. I Nordrhein-Westfalen är *Gymnasiale Oberstufe* kursbaserad och indelad i en introduktionsfas (första året) och en kvalifikationsfas (år två och tre). I introduktionsfasen är betygsskalan densamma som i grundskolan, 1–6. I kvalifikationsfasen förekommer även ett system där 15 betygspoängsteg ersätter 1–6-skalan (se tabell 2).

Den som har godkända lärarsatta betyg och klarar *Abitur*-examensproven i slutet av sin utbildning kan söka till universitet som antar

¹¹² IQB (2021a).

¹¹³ IQB (2021a).

¹¹⁴ IQB (2021b).

¹¹⁵ IQBa.

huvudsakligen baserat på genomsnittsbetyg vid konkurrens om platser. *Abitur*-proven i sig ses generellt som behörighetsgivande men det finns vanligen ingen ämnesbehörighetskoppling mellan examensprovsämnen och den utbildning som söks. Det är dock vanligt förekommande att utbildningar kräver särskilda antagningsprov för alla eller en viss kvotgrupp av sökande. En särskild egenhet vid tyska lärosäten är tillämpning av en egen kvotgrupp för sökande som har haft uppehåll mellan gymnasiestudierna och högskolan. Ju längre uppehåll, desto större chans att få en plats.¹¹⁶

Abitur-proven

Sedan 2012 finns en nationell standard för den tyska studentexamen, *Abitur*.¹¹⁷ Alla utom en delstat, där skolorna själva konstruerar sina prov, har centraliserat studentexamen i någon utsträckning. IQB ansvarar för utveckling av en årlig bank med provfrågor som delstaterna kan använda i ämnena tyska, engelska, franska, matematik, biologi, kemi och fysik. Proven prövas inte ut, utvärderas inte psykometriskt och det sätts ingen godkäntgräns från centralt håll.

Varje delstat bestämmer själv vilka frågor de använder och om de vill komplettera med egna frågor. Varje delstat uppmanas att välja 50 % av sina provuppgifter från provbanken och att inte ändra i konstruktion eller formulering. Ett typiskt prov i ett ämne innebär att eleven ska välja en av fyra möjliga uppgifter och besvara den på 300 minuter.

I Nordrhein-Westfalen beslutas alla prov centralt utom de muntliga, vilka utformas av enskilda lärare och blir externt granskade innan de genomförs.¹¹⁸ En elev gör vanligen fyra prov varav tre skriftliga och ett muntligt. Eleven ska välja minst ett provämne från vardera av följande tre ämneskategorier: matematik/NO-ämnen, språk och SO-ämnen. I Nordrhein-Westfalen läser elever vanligen två ämnen på avancerad nivå på gymnasial nivå och skriver två av sina prov i just dessa, plus ytterligare två eller tre, ämnen.

IQB förser delstaterna med anvisningar för bedömning av proven inklusive förslag på vilka typer av svar som bör motsvara ett visst

¹¹⁶ Se t.ex. Technische Universität Berlin, <https://www.tu.berlin/en/studierendensekretariat/topics-a-z/waiting-periods>.

¹¹⁷ IQBb.

¹¹⁸ Ministerium für Schule und Bildung des Landes Nordrhein-Westfalen (2024).

provbetyg. Varje prov bedöms av två lärare, en från den egna skolan och en från en annan skola. Den som missar ett prov eller blir underkänd ges en eller två nya möjligheter i nära anslutning till ordinarie provtillfälle. Det förekommer också att den som, av giltigt skäl, missar ett centralt framtaget prov erbjuds att i stället skriva ett lärarkonstruerat prov. Elever som inte klarar *Abitur*-examen, får gå om kvalifikationsfasens andra år.

I slutbetyget kan en elev få maximalt 900 betygspoäng och 500 betraktas som godkänt. Examensproven kan ge maximalt 300 poäng medan de lärarsatta betygen från elevens två sista år på *Oberstufe* ger upp till 600 poäng. Varje moment som ingår måste bidra med minst en poäng. Poängen konverteras till ett övergripande betyg på 1–6-skalan, vilken används vid antagning till högskolan.

Alberta (Kanada)¹¹⁹

Skol- och betygssystem

I Kanada finns inget federalt utbildningsdepartement och inget nationellt utbildningssystem. Ansvar för utbildning vilar på de 13 enskilda provinserna och territorierna i enlighet med Kanadas konstitution. Varje provins och territorium har lagstiftning, policyer, program, läroplaner och praxis för grundskole- och gymnasieutbildning. Likaså varierar åldersintervallet för obligatorisk skolgång och betygssystemens konstruktion.

Alberta har kallats ”*the jurisdiction in the English-speaking world with the best state school performance*”.¹²⁰ Bland de kanadensiska provinserna och territorierna uppvisar Alberta ofta bäst resultat i internationella kunskapsmätningar.¹²¹ Resultaten i PISA 2022 var inget undantag; Alberta överträffades endast av Singapore. Därtill är lärarna i Alberta bland de nöjdaste, med löneläget och läraryrkets status i samhället, i OECD.¹²²

¹¹⁹ Huvudsaklig källa för avsnittet: *Guide to Education*, Government of Alberta, <https://www.alberta.ca/guide-to-education>.

¹²⁰ Gove (2010, s. 20).

¹²¹ The Alberta Teachers' Association (2023).

¹²² Education GPS (u.å.).

Första delen av grundskolan i Alberta, *elementary school*, är sexårig och påbörjas vid sex års ålder. Därefter följer sex år på *high school* (tre år vardera på junior respektive senior nivå).¹²³

Enligt Albertas skollag ska lärare kontinuerligt bedöma elevernas kunskapsutveckling och rapportera resultaten till elever, vårdnadshavare och skolans styrelse. Lärarna är också skyldiga att delta i provinsens arbete med utveckling och rättning av centrala prov.¹²⁴ Det framgår dock inte av de officiella dokument som vi har fått tillgång till, eller i våra kontakter, vid vilken ålder eller i vilka årskurser som lärarsatta betyg ska sättas. Inte heller framgår någon lagstadgad betygsskala eller enhetligt system för vare sig Alberta eller Kanada som helhet. Det tycks förekomma bokstavsskalor, numeriska skalor, godkänt/icke godkänt och bedömning i procent.¹²⁵ Varje skolhuvudman ansvarar för att ha en plan för arbetet med bedömning inklusive riktlinjer för sådant som elevers överklagandemöjligheter och återkoppling till elever och lärare.¹²⁶

Centrala prov i grundskolan

Från och med läsåret 2022/23 är det obligatoriskt att genomföra diagnostiska tester i läsning och matematik i årskurs 1 till 3 i *elementary school*. Elever vars kunskaper identifieras ligga i riskzonen ska genomgå uppföljande bedömning i slutet av läsåret. Skolor väljer själva tester från en centralt framtagna lista.¹²⁷ Fokuset på formativ bedömning är starkt i Alberta. Det har exempelvis gjorts försök med ytterligare obligatoriska centrala digitala prov (*SLA, Student Learning Assessments*) i språk och matematik vid början av årskurs 3 som ytterligare ett sätt att identifiera elevers behov. Dessa har dock fått kritik från bl.a. lärare för att vara administrativt betungande och stressande för eleverna; de har därför pausats.¹²⁸

¹²³ Government of Alberta (2024a).

¹²⁴ Education Act, Statutes of Alberta, 2012 Chapter E-0.3, Section 196.

¹²⁵ Nandhini (2024).

¹²⁶ Council of Ministers of Education, Canada (2023).

¹²⁷ Government of Alberta (2023a).

¹²⁸ Government of Alberta (2024b) och Konguavi (2019).

Provinsgemensamma prov, *Provincial Achievement Testing (PAT)*, görs i årskurs 6 och 9 i engelska, franska, matematik, SO- och NO-ämnen. Syftena med proven är att

- följa upp om eleverna lär sig vad de förväntas lära sig,
- informera befolkningen i Alberta om hur väl elever uppnår provinsens förväntade kunskapsnivå vid givna punkter i sin skolgång, samt
- hjälpa skolor, skolmyndigheter och provinsen att följa upp och förbättra elevernas lärande.¹²⁹

Huvuddelen av proven utgörs av skriftliga flervalfrågor men det förekommer även skrivuppgifter i matematik och språk. Extra provtid eller särskilda hjälpmedel kan medges för elever med särskilda behov. En övergång till digitala prov har påbörjats läsåret 2023/2024.¹³⁰ I syfte att främja jämförelser över tid används ankaruppgifter och kravgränssättningen för vad som anses beteckna *Acceptable Standard* och *Standard of Excellence* sätts i efterhand.¹³¹ På så vis bibehålls andelarna elever som får respektive vitsord över tid. Proven är sekretessbelagda även efter att de skrivits.¹³² Vissa provdelar bedöms först lokalt av elevernas egna lärare och därefter centralt (av lärare som anmäler intresse för att vara centrala bedömare) medan andra delar endast bedöms centralt.¹³³ I varje provämne ska betyg sättas på bokstavs- eller procentskala enligt följande: A = 80–100 %, B = 65–79 %, C = 50–64 %, F = < 50 %.

Pan-Canadian Assessment Program (PCAP) är en nationell kunskapsmätning som sedan 2007 genomförs vart tredje år av ett urval av elever i årskurs 8.¹³⁴ Proven omfattar alltid läsning, matematik och naturvetenskap, men fokusområdet varierar. PCAP är inte kopplad till individens betyg eller utbildningsmöjligheter utan syftar till att komplettera enskilda provinsers uppföljningar och ge nationellt jämförbar information om kunskapsnivåerna hos elever i årskurs 8.

¹²⁹ Government of Alberta (2023b).

¹³⁰ Government of Alberta (2024c).

¹³¹ Government of Alberta (2023c).

¹³² Government of Alberta (2023b, s. 11).

¹³³ Government of Alberta (2023b).

¹³⁴ Council of Ministers of Education, Canada (u.å.).

Examensprov för *High School Diploma*

Behörighet till högre utbildning bygger framför allt på *High School Diploma*-examen i årskurs 12 som kräver att eleven har slutfört studier i vissa ämnen, inklusive idrott och *Career and Life Management*, samt genomförda examensprov i de ämnen som lästs på avancerad nivå. Examensprov kan göras i engelska, franska, matematik, biologi, kemi, fysik och SO och eleven ska ha läst klart och fått lärarsatta betyg i kurser i provämnena.¹³⁵ Varje prov skrivs på specifika tider och dagar och de är oftast sekretessbelagda efter att de skrivits. Under år 2024 erbjuds fem provtillfällen i varje ämne och de ges i form av skriftliga prov som kan skrivas digitalt eller på papper. Den digitala provplattformen är ny fr.o.m. läsåret 2024/25 och är valfri för skolor att använda.¹³⁶ Den erbjuder möjlighet för elever att träna för att bli bekanta med plattformens utseende och hur provfrågor kan se ut.

Proven skrivs på skolorna, men lärare som undervisat i provämnena får inte närvara under proven. Extra provtid eller särskilda hjälpmedel kan medges för elever med särskilda behov. Undantag från att helt skriva examensproven kan beviljas av Albertas centrala skolmyndighet, dvs. inte på skolnivå. Den som missar ett enskilt delprov kan få ett provbetyg framräknat utifrån det lärarsatta betyget och betyget på de delprov som skrivits.

Vissa prov rättas maskinellt men skrivuppgifter bedöms centralt av två lärare som nomineras av sina skolor till att vara bedömare. Resultat och provbetyg redovisas som en procentsats. I syfte att betygen ska vara jämförbara mellan skolor och över tid används ankaruppgifter och kravgränssättning i efterhand (*equating*).¹³⁷

Slutbetyget i ämnen som eleven läst på avancerad nivå utgörs av en sammanvägning av lärarsatta betyg (70 %) och provresultat (30 %). Minst 50 % krävs för godkänd examen. Både provbetygssättning och lärarbetygssättning kan överklagas. Den som blir underkänd och behöver skriva om ett eller flera examensprov kan registrera sig och får betala en avgift för det.

¹³⁵ Government of Alberta (2023d).

¹³⁶ <https://abed.vretta.com/#/en/login-router-st>.

¹³⁷ Government of Alberta (2023e).

Antagning till högre utbildning

Kanadensiska universitet har skilda behörighetskrav och urvalsgrunder för antagning och kraven kan variera beroende på om den sökande har sin gymnasieutbildning från samma provins som universitetet ligger i eller ej. Genomsnittsbetyg spelar normalt sett stor roll som urvalsprincip när det är fler sökande än utbildningsplatser, men det förekommer också antagningsprov och alternativ antagning för elever med särskilda meriter (t.ex. inom idrott, kultur eller teoretiska studier) eller bakgrund i Kanadas ursprungsbefolkning.

På University of Alberta ställs följande minimikrav på betyg för beräkning av genomsnitt som används vid ansökan: betyg i varje ämne i årskurs 11: 60 % eller högre; betyg i varje ämne i årskurs 12: 50 % eller högre; 70 % som lägsta genomsnitt i de fem obligatoriska ämnen som krävs för behörighet till det program som söks.¹³⁸ Även övriga större lärosäten i Alberta tillämpar liknande urvalsgrunder.¹³⁹ Många yrkesinriktade programutbildningar, såsom jurist, tandläkare eller dietist, är organiserade som påbyggnadsutbildningar. De kräver att studenterna först läser vissa grundläggande fristående kurser eller tar en kandidatexamen för att sedan kunna antas till påbyggnaden baserat på de inledande studieresultaten tillsammans med andra grunder såsom personligt brev, rekommendationer och intervjuer.¹⁴⁰

Hongkong¹⁴¹

Skol- och betygssystem

Förskolan är inte obligatorisk i Hongkong. Grundskolan, som påbörjas vid sex års ålder, är uppdelad i *primary school* och *secondary school* och omfattar totalt 12 års undervisning. Tidigare skedde en uppdelning baserad på *high stakes*-prov även efter *junior secondary school*.¹⁴² De är avskaffade men den som vill kan avsluta *secondary school* efter halva tiden och gå vidare till en treårig yrkesskola. Den

¹³⁸ University of Alberta (2024).

¹³⁹ Se t.ex. University of Lethbridge, <https://www.ulethbridge.ca/future-student/canadian-high-school-admission-requirements>, och University of Calgary, <https://www.ucalgary.ca/future-students/undergraduate/requirements>.

¹⁴⁰ University of Alberta (2024).

¹⁴¹ Huvudsakliga källor för avsnittet: e-postkontakt med Education Bureau och Hong Kong Examinations and Assessment Authority.

¹⁴² Creese m.fl. (2016).

som fortsätter till senior nivå studerar de obligatoriska ämnena kinesiska, engelska, matematik och *liberal studies* samt upp till tre valfria andra ämnen som kan vara SO- och NO-ämnena, språk eller något mer professions-/yrkesinriktat ämne. Den högskoleförberedande gymnasieskolan avslutas med examensprov som ger ett *Hong Kong Diploma of Secondary Education (HKDSE)*.

Antagning och urval till *secondary school* sker i huvudsak centraliserat och baseras på en kombination av sökandes önskemål och deras tidigare studieresultat. Elevernas resultat på *Internal Assessments (IA)* används som mått och en standardisering jämfört med samma skolas tidigare årskohorter görs inför en Hongkong-gemensam skalning.¹⁴³ Skolorna får därutöver själva sköta antagningen till högst 30 % av platserna. Det innebär att de kan tillämpa egna antagningskrav för dessa.¹⁴⁴

Hongkong har en centralt organiserad resursfördelning till de offentligtfinansierade skolorna som bl.a. innebär en särskild strategi för att stödja lågpresterande elever i *secondary schools*. I stället för justerad styrning genom finansiering ökas antalet lärartjänster på skolor med ett stort antal lågpresterande elever. Omfattningen av elever med behov skattas baserat på resultaten på interna skolutvärderingar som görs under de två sista åren av *primary school* och under antagningsprocessen till *secondary school*. Elevbedömningarna placeras på en gemensam skala och därefter beräknas hur många extra lärare varje skola beräknas behöva utifrån sitt elevunderlag.¹⁴⁵

TSA-proven¹⁴⁶

Territory-wide System Assessment (TSA) är centralt framtagna skriftliga prov som görs av alla elever i kinesiska, engelska och matematik i motsvarande de svenska årskurserna 3, 6 och 9 (dvs. årskurs 3 och 6 i *primary school* och årskurs 3 i *junior secondary school*). *TSA Moderation Committees* arbetar med provkonstruktion och kvalitetssäkring. Inga prov är digitala. Bedömning görs centralt och anonymt.

Den ansvariga skolmyndigheten för TSA betonar att proven inte kräver förberedelse och har ett huvudsakligen diagnostiskt syfte.

¹⁴³ Yung (2020).

¹⁴⁴ Education Bureau (2023) och e-post med Andrew Chung.

¹⁴⁵ National Center on Education and the Economy (2024a).

¹⁴⁶ Hong Kong Examinations and Assessment Authority (u.å.).

Det har dock riktats kritik mot systemet eftersom enskilda skolor tycks kräva omfattande förberedelser och extraarbete av eleverna i syfte att maximera resultaten. Det har uttryckts önskemål om att åtminstone ta bort proven för de yngsta eleverna. Resultaten återkopplas till skolorna på skolnivå, inte individnivå, och dokumenteras därför inte i elevens resultatdokument eller används för ansökan eller kvalifikation till högre årskurs.

HKDSE-proven

Alla prov i de 24 så kallade kategori-A-ämnena, som är de vanligaste ämnena som stora grupper studenter gör prov i, görs på papper och bedömningen görs normalt sett centralt och anonymt. Proven bedöms av två olika personer. Om deras bedömning skiljer sig mycket åt kan en tredje och fjärde bedömning göras. Principen vid oenighet är att de två bedömningar som ligger närmast varandra används.¹⁴⁷ Standarden för betygsgränser för varje enskilt prov och provtillfälle sätts i efterhand. Det görs genom att en panel av experter granskar ett antal exempel på vad som ska utgöra välskrivna prov och jämför resultaten med fjolårets i samma ämne. De har också statistiska metoder till sin hjälp. Baserat på deras granskning rekommenderas gränserna för varje provdel i varje provämne.

Flera insatser i syfte att säkerställa provstabilitet har vidtagits. Genomförda provfrågor analyseras och nya provfrågekonstruktioner prövas ut för kommande prov.¹⁴⁸ Analyser av materialet används också för att avgöra vilka gränser som ska gälla för olika betygsnivåer så att dessa hålls stabila över tid. Provpöängen räknas om till betyg (tabell 3).

Tabell 3 Provpöäng och betygsbetygsskalor för Hongkongs HKDSE-prov

5**	A*	Top of the class
5*	A/A*	Excellent
5	A	Good
4	B/C	Adequate
3	D/E	Pass
2–1	F	Fail

Källa: Hong Kong Examinations and Assessment Authority (2018).

¹⁴⁷ Hong Kong Examinations and Assessment Authority (2021).

¹⁴⁸ Hong Kong Examinations and Assessment Authority (2018).

En tumregel är att de 10 % bäst presterande av eleverna som får 5 får 5** och de därpå följande 30 % får 5*, men variationer förekommer. I tolv kategori-A-ämnen gör eleverna också provdelar som bedöms av deras lärare, *School-Based Assessments, SBA*. Det kan vara muntliga, skriftliga eller laborativa uppgifter enskilt eller i grupp. Skolor skickar in exempel på elevarbeten och de inrapporterade resultaten modereras centralt. Modereringen kan ske både statistiskt och genom expertbedömning.¹⁴⁹ Med statistisk moderering menas att skolans genomsnittliga lärarsatta betyg i ett givet ämne (skolans genomsnittliga SBA) kalibreras till att motsvara skolans genomsnittliga resultat på det nationella provet (*public examination*) i det ämnet. Även spridningen i skolans lärarsatta betyg (SBA) kalibreras till viss del baserat på motsvarande spridning i det nationella provet. Rangordningen av elever inom en viss skola förändras dock inte av den statistiska modereringen.¹⁵⁰ Betygssättningen kan inte överklagas. I övriga ämnen, som färre studenter skriver prov i, används andra bedömningsätt och betygsskalor, exempelvis *unattained/attained/attained with distinction* eller a–e.

I diplommet listas betygen i respektive ämne, ibland på delprovsnivå. Inget numeriskt betygsgenomsnitt räknas ut. Antagningen till högre utbildning är decentraliserad och lärosätena har egna antagningsprinciper och kriterier för bedömning av ansökningar. Eftergymnasiala utbildningsinstitutioner ska sträva efter att säkerställa att endast de mest lämpade och välförtjänta sökande antas och de uppges ofta göra en samlad bedömning av betyg, antagningsprov, intervju prestationer, personliga egenskaper, icke-akademiska prestationer, intressen, erfarenheter och sökandes programpreferenser. Gymnasieelever som vill söka till universitet i fastlands-Kina behöver dock vanligen göra särskilda kinesiska antagningsprov, *gaokao*, för att konkurrera om platserna där.

¹⁴⁹ Hong Kong Examinations and Assessment Authority (2023a).

¹⁵⁰ Hong Kong Examinations and Assessment Authority (2023b).

Singapore¹⁵¹

Skol- och betygssystem

Singapore strävar efter att kombinera hög likvärdighet och utbildningskvalitet med världsledande prestationer och utmärker sig bland OECD-länderna genom sin låga andel lågpresterande elever och höga andel högpresterande elever, inklusive en mycket hög andel högpresterande elever med låg socioekonomisk status. Totalt bedöms tio procent av eleverna från resurssvaga hushåll vara högpresterande i Singapore, jämfört med ett OECD-genomsnitt på tre procent.¹⁵²

Det starkt centraliserade skolsystemet skiljer sig i många avseenden från andra länders system. Utbildningsdepartementet fördelar medel till alla skolor, fastställer kursplaner och nationella prov, övervakar lärarlegitimationer, hanterar utvärderings- och befordrings-systemet för skolpersonal samt anställer rektorer och lärare. Departementet ger utmärkelser till skolor med särskilt goda resultat och kan ge utmärkelser och bonus på lönen till lärare.

Alla elever ges ett så kallat *Edusave*-konto till vilket regeringen bidrar med medel, cirka 2 000 kronor per år, men resurssvaga hushåll får mer än andra. Bidraget kan användas för skolavgifter, som är mycket låga i den offentligfinansierade skolan, och andra utbildningskostnader. Vårdnadshavarna väljer hur medlen används.¹⁵³ Elever som uppfyller vissa krav, såsom goda skolprestationer, anmärkningsvärd förbättring eller utmärker sig vad gäller ordning och uppförande, kan få stipendier som ger ytterligare medel.¹⁵⁴

Förskolan är inte obligatorisk i Singapore. Grundskolan, som påbörjas vid sex års ålder, är uppdelad på sex år i *primary school* och fyra eller sex år i *secondary school*. Den första screeningen av barns läs- och räkneförmåga äger rum den första månaden i grundskolan. Skolbaserade prov och bedömning i de lägre årskurserna är dock huvudsakligen diagnostiska och formativa till sin karaktär. Från och med årskurs 5 kan eleverna vanligen välja olika nivåer i vissa ämnen.

Bedömning av elevers prestationer under grundskolan återkopplas regelbundet till vårdnadshavarna i formuläret *Holistic Development*

¹⁵¹ Huvudsakliga källor för avsnittet: hemsidorna för Ministry of Education Singapore, <https://www.moe.gov.sg>, och Singapore Examinations and Assessment Board, <https://www.seab.gov.sg>.

¹⁵² National Center on Education and the Economy (2024b).

¹⁵³ Ministry of Education Singapore (2024a).

¹⁵⁴ Ministry of Education Singapore (2024b).

Profile (HDP). I de första årskurserna används ingen betygsskala för skolämnena utan fokus är på kvalitativa omdömen av elevernas lärande och utveckling. Däremot redovisas elevens betyg i uppförande, omdöme om personliga egenskaper, en redogörelse för elevens engagemang i fritidsaktiviteter och skolnärvaro samt en bedömning av elevens fysiska kondition, bl.a. i form av *body mass index*.¹⁵⁵

Från och med årskurs 3 innehåller HDP *marks* (0–100) och resultat av helhetsbedömningar i varje ämne, s.k. *weighted assessments*. Med start 2019 har förändringar gjorts i syfte att minska fokus på akademiska prestationer och stärka elevernas holistiska utveckling och engagemang. Numera har eleverna därför inte regelrätta examinationer eller får betyg i de två första årskurserna och decimalerna vid *mark*-redovisningen samt elevens rankning i sin klass och klassens medelvärde har tagits bort från HDP.¹⁵⁶ Därtill har *mid-year exams* tagits bort från tre högre årskurser.

PSLE-proven

I slutet av *primary school* gör alla elever *Primary School Leaving Examination* (PSLE) i fyra ämnen: engelska, matematik, naturvetenskap och modersmål. Eleverna gör proven på den nivå som de läste ämnet på i årskurs 5 och 6. PSLE-proven administreras av *Singapore Examinations and Assessment Board* och det finns digitala, skriftliga och muntliga provdelar.¹⁵⁷ Paneler med ämnesspecialister och tidigare examinatoreer anlitas för att bidra till konstruktionen av proven, nå en lagom svårighetsgrad på frågorna och ta fram preliminära bedömningsriktlinjer.¹⁵⁸ När eleverna har skrivit proven standardiseras och uppdateras bedömningsriktlinjerna så att bedömningarna blir fullt jämförbara med tidigare provomgångar. En bredd av elevsvar samlas in och studeras och en grupp, av företrädesvis seniora rektorer, utses till *Key Marking Personnel* (KMP) för att säkerställa kvaliteten i bedömningsriktlinjerna. Bedömningen sker centralt under ledning av en chefsbedömare och KMP gör stickprovsbedömning.

Elever kan köpa tidigare års provversioner att träna på och elever med särskilda behov kan få extra provtid, undantas från provdelar

¹⁵⁵ Ministry of Education Singapore (2018a).

¹⁵⁶ Ministry of Education Singapore (2018b).

¹⁵⁷ Singapore Examinations and Assessment Board (2024a).

¹⁵⁸ Singapore Examinations and Assessment Board (2021).

eller få använda hjälpmedel såsom ordböcker.¹⁵⁹ Regelverket medger särskilda arrangemang för att skriva proven för elever som t.ex. är på sjukhus eller har smittsamma sjukdomar. Bedömningen av proven kan överklagas mot en avgift.

Användningen av PSLE-resultaten har förändrats över tid. Sedan år 2021 betygssätts inte elevernas individuella ämnesprestationer utan elevens sammantagna resultat jämförs med andra elevers för ranking. PSLE-poängen härleds från summan av ämnespoängen, som varierar från 4 till 32. Poängen på proven konverteras till en individuell placering på åtta *Achievement Levels* (1 är bäst, 8 är sämst), som används för att avgöra vald linje för den kommande lägre sekundärutbildningen, liksom vilken skola eleven kan gå på. Det går att söka till sex skolor och skolorna antar till stor del baserat på sökandes PSLE-ranking. Det är dock tillåtet med viss direktantagning, dvs. att ta in elever baserat på talang inom vissa ämnen, utan att ta hänsyn till PSLE-resultat, för att skapa större mångfald. Utbildningsministeriet hjälper till att placera de elever som inte kommer in på någon skola de valt.

Linjer och GCE-examensprov

När eleverna börjar *secondary school* väljer de mellan tre linjer: Express, Normal (akademisk) och Normal (teknisk). Alla linjer erbjuder samma studiegång, men Express är snabbare och Normal (teknisk) erbjuder fler praktiska moment. I de flesta fall är det elevernas resultat på PSLE-proven som avgör på vilken linje de kommer att gå, men föräldrar och elever kan förespråka olika linjer om de visar prov på snabbare inlärning eller behöver mer hjälp. Singapore inför för närvarande också ett system där eleverna kan välja olika inriktning i olika ämnen, snarare än en övergripande linje som gäller för samtliga elevens ämnen. En elev kan t.ex. välja en teknisk inriktning i matematik, men en expressinriktning i engelska. Denna förändring, *subject based banding*, *SBB*, är redan praxis i alla *primary schools* och har införts i alla *secondary schools* under 2024.¹⁶⁰

Gymnasieskolan avslutas med examensprov, *General Certificate of Education (GCE) examinations*. GCE-proven bedöms centralt

¹⁵⁹ Singapore Examinations and Assessment Board (2024b).

¹⁶⁰ Kai & Shafeeq (2023).

med stöd av Cambridge Assessment International Education. Efter fyra års studier avlägger eleverna prov på avancerad nivå eller *ordinary* nivå på expresslinje och prov på normalnivå på normal (teknisk) linje. Elever på normal (akademisk) linje kan avlägga normalnivåprov efter fyra års studier eller *ordinary*-nivåprov efter fem år. Den som är godkänd på prov i minst ett provämne får ett examensbevis. Antal prov och betygsskalor varierar enligt följande:

- GCE Normal teknisk, prov i 5–7 ämnen, betygsskala A–E (D = godkänt)
- GCE Normal akademisk, prov i 5–8 ämnen, betygsskala 1–6 (6 = underkänt och 1 = högst)
- GCE Express *Ordinary*, prov i 6–9 ämnen, betygsskala A1, A2, B3, B3, B4, C5, C6, D7, E8, 9. (9 = underkänt och minst C6 krävs i ett ämne för examensbevis)
- GCE Express Avancerad, prov i 7–10 ämnen, betygsskala A–E/*subpass/ungraded* i ämnen på högre nivå, H1 och H2; betygsskala *Distinction, Merit, Pass, Ungraded* i ämnen på lägre nivå H3 och vissa muntliga prov.¹⁶¹

Från och med 2027 kommer normalnivå- och *ordinary*-nivåprov att ersättas av *Singapore-Cambridge Secondary Education Certificate (SEC)*.¹⁶² De nya proven avser att bättre avspegla de nivåer som eleverna läst respektive ämne på. Elever kommer också att ges möjlighet att gå ytterligare ett år på gymnasiet för att läsa mer krävande kurser.

Antagning till högre utbildning

Antagning till högre utbildning beror på vilken gymnasieutbildning den sökande har samt resultaten på GCE-proven. Singapore har sex universitet till vilka studenter med goda resultat från examensproven på GCE Express avancerad nivå är behöriga att söka.¹⁶³ Enskilda utbildningar kan därutöver ha särskilda behörighetskrav och antagningsprov. För den som inte uppfyller behörighetskraven finns

¹⁶¹ Universities and Colleges Admissions Service (2019).

¹⁶² Ministry of Education Singapore (2024c).

¹⁶³ Ministry of Education Singapore (2023).

också eftergymnasial extrautbildning. Detta erbjuder ytterligare en väg till framtida möjligheter till universitetsutbildning och yrkesutbildning.¹⁶⁴

Endast ett lärosäte, National Institute of Education (NIE) på Nanyang Technological University, utbildar lärare. På så sätt begränsar Singapore lärarekryteringen till de studenter som är kvalificerade för studier på landets forskningsuniversitet. Årligen beräknas hur många nya lärare som behövs och sen erbjuds just så många utbildningsplatser. Studenterna på lärarutbildningarna får stipendium och i genomsnitt antas endast en av åtta sökande. Antagningen baseras på de sökandes examensprovsbetyg men inkluderar också intervjuer och bedömning av den sökandes tidigare akademiska meriter och bidrag till skolan och samhället.¹⁶⁵

New South Wales (Australien)¹⁶⁶

Skol- och betygssystem

Den obligatoriska grundskolan i New South Wales inleds med ett års förskola (*kindergarten*). Därefter följer sex år i *primary school*, som börjar vid sju års ålder. *High school*, med årskurserna 7–12, erbjuder motsvarande högstadie- och gymnasieutbildning. I årskurs 10 väljer eleverna vilka ämnen (normalt 5–6 ämnen) de ska fokusera på under de sista två åren (11–12), vilket normalt sett är de ämnen som de sen gör examensprov i och har i sin gymnasieexamen (*Higher School Certificate, HSC*). Valet kan även omfatta yrkesinriktade kurser och vissa av dessa har frivilliga examensprov i slutet av gymnasieskolan. Elever som redan tidigare vet att de vill gå en yrkesutbildning kan välja detta redan tidigare under *high school*.¹⁶⁷ Skriftliga omdömen ges varje termin eller läsår fr.o.m. förskolan, men dessa har ingen formell betydelse för fortsatta studier.¹⁶⁸ Med början i årskurs 1 ska omdömena innehålla lärarsatta bokstavs-betyg.

¹⁶⁴ Ministry of Education Singapore (2024d).

¹⁶⁵ National Center on Education and the Economy (2024b).

¹⁶⁶ Huvudsakliga källor för avsnittet: muntlig information från New South Wales Education Standards Authority och NSW Education Standards Authoritys hemsida, <https://www.nsw.gov.au/education-and-training/nesa>.

¹⁶⁷ New South Wales government (u.å.).

¹⁶⁸ New South Wales government (2023).

Den skolbaserade betygssättningen i den elvaåriga grundskolan görs vanligen på en femgradig betygsskala som går från A till E där A är det högsta betyget.¹⁶⁹ Det finns beskrivningar för varje betygssteg. Till och med 2011 fanns examensprov i årskurs 10, men de är nu avskaffade. Elever som går i *Support Classes* eller *Special Support Schools*, motsvarande Sveriges resursskola och anpassad grundskola, kan även ges betyget UG (*Ungraded*).¹⁷⁰ Detta betygssteg kunde fram till 2020 även användas i andra skolformer men har därefter reserverats för elever som inte kan följa ordinarie skolgång.

Nationella prov

National Assessment Program – Literacy and Numeracy (NAPLAN) är formativa, nationella (för hela Australien), digitala prov som görs årligen av elever i årskurs 3, 5, 7 och 9.¹⁷¹ Proven ska göra det möjligt för vårdnadshavare och lärare att se hur elevernas läs- och räknefärdigheter utvecklas över tid. Resultaten presenteras såväl individuellt som på skol-, delstats- och nationell nivå.

Personer med expertis inom läroplaner, ämnesinnehåll, bedömning och psykometri arbetar med framtagningen. Läroplans- och bedömningsexperter från varje delstat och territorium, samt representanter från den icke-statliga sektorn, granskar föreslagna provuppgifter för att säkerställa att de uppfyller kraven. Representativa urval av elever från varje stat och territorium deltar i utprovningen av provfrågor. Resultaten analyseras av oberoende psykometriska experter.

Flervalsfrågor rättas automatiskt medan skrivuppgifter bedöms av utbildade, externa bedömare. Varje stat och territorium ansvarar för sin rättning och bedömning. När proven har skrivits och bedömts ekvivaleras de med tidigare provutgåvor så att resultat kan rapporteras på samma mätskalor oberoende av provens potentiellt varierande svårighetsgrad. Beroende på antal poäng och aktuellt års poänggränser kan eleven få följande omdömen på vart och ett av de fem delproven: över förväntan, stark, under utveckling och behöver ytterligare stöd. Elever får inget sammanfattande provbetyg men i återkopplingen till vårdnadshavarna redovisas den enskildes resultat

¹⁶⁹ New South Wales Education Standards Authority (u.å.).

¹⁷⁰ ACARA (2023).

¹⁷¹ ACARA (2024).

parallellt med det genomsnittliga resultatet för alla elever i samma årskurs i Australien.

Alla elever skriver inte proven vid exakt samma tillfälle och i syfte att förhindra spridning av provuppgifter används adaptiv teknik som innebär att elever automatiskt får olika följdfrågor baserat på hur de besvarar inledande frågor. Uppgifterna ska matcha elevernas prestationsnivå.

Vidare finns obligatoriska tester, *Minimum Standards Tests for Literacy and Numeracy*, för elever i årskurserna 10 och uppåt i New South Wales. Dessa syftar till att säkerställa att elever uppnår delstatens mål om en förväntad miniminivå i matematik och läs- och skrivkunnighet. Individerna ska visa att den kan behärska vardags-situationer såsom att följa säkerhetsanvisningar i utrustningsmanu-aler, förstå information om mobilabonnemang, skriva en jobbansökan samt göra veckobudget.¹⁷² Eleverna får göra om testerna tills de har klarat dem med godkänt resultat.

Betyg och bedömning i gymnasieskolan

I *high school* sätter lärarna bokstavs-betyg i alla ämnen i en *Record of School Achievement (RoSA)*. RoSA är en elevs studiedokumentation som t.ex. kan användas av elever som inte fullföljer gymnasieskolan och får en examen.¹⁷³ År 12 används i stället sifferbetyg som dokumenteras i en *Higher School Certificate Record of Achievement*.¹⁷⁴ För kurserna som ingår i HSC finns beskrivningar av olika uppgifter och bedömningskriterier, hur tungt de olika uppgifterna ska viktas, när de ska genomföras och hur de ska betygsättas. För att säkerställa att betygen speglar elevernas kunskaper i ämnet på en inom delstaten jämförbar nivå skickas stickprov av elevens betygsunderlag vidare till *HSC Record of Achievement*.¹⁷⁵ De olika stickproven utvärderas sedan och ett betyg sätts för varje elev. Stickproven fungerar som en typ av social moderering som kan vägleda lärarnas bedömning och fungera som fortbildning av deras bedömarkompetens.¹⁷⁶ Efter genomförda uppgifter ges eleverna återkoppling så att de kan förbättra sina

¹⁷² New South Wales government (2024a).

¹⁷³ New South Wales government (2024b).

¹⁷⁴ New South Wales government (2024a).

¹⁷⁵ New South Wales Education Standards Authority (u.å.).

¹⁷⁶ New South Wales Education Standards Authority (u.å.).

kunskaper. Eleverna ges däremot inte tillgång till betygsresultatet, utan enbart sin rankning inom skolan. Det beror på att de lokala, lärarsatta betygen modereras för att länka dem till uppnåendemålen.¹⁷⁷ Lärare och skolor uppmanas således att inte lämna information till eleverna om vilka betyg de satt utan anmoda dem att vänta tills examensprov och moderering är klara.

HSC-examensproven¹⁷⁸

Examensproven som ligger till grund för HSC-examen ges en gång per år och syftar till att pröva hur väl eleverna har uppnått kunskapsmålen (*learning outcomes*) i kursplanerna samt för att moderera de lärarsatta betygen på skolnivå. Det finns beskrivningar för vad eleverna ska kunna för de olika betygsstegen. Proven är konstruerade för att pröva ett representativt urval av kunskap, förståelse och förmågor i syfte att undvika att proven blir förutsägbara. De ska i stället uppmuntra eleverna att förbereda sig mot hela kursplanens betygsriterier (*performance band descriptions*). Över tid förväntas proven pröva samtliga kursplanemål (*construct coverage*). Proven utformas även för att pröva eleverna mot olika kunskapsnivåer och även avancerade förmågor (*high order skills*).

Examensproven bedöms centralt av tränade bedömarlag för att säkerställa att bedömningarna blir likvärdiga (interbedömarreliabilitet). Bedömningsarbetet leds av seniora bedömare som utformar riktlinjer som beskriver vad som krävs för de olika betygsnivåerna och elevexempel (*benchmarks*) tas fram för att kalibrera mot. Bedömarlagen får träna på att bedöma utifrån dessa riktlinjer till dess att de seniora bedömare anser att bedömningsanvisningarna tillämpas korrekt. Under det fortsatta bedömningsarbetet övervakas bedömarlagen av de seniora bedömare så att avvikelser inte ska uppstå. Dessa bestämmer även kravgränserna för de olika betygen. Efter bedömningen sammanfattas en råpoäng för provet. För vissa prov räknas råpoäng om till en standardiserad 100-poängsskala.

I en elevs examen kombineras de lokala lärarsatta betygen med examensprovsbetygen. För att betygen ska bli jämförbara mellan skolor modereras betygen. Modereringen sker, på liknande sätt som

¹⁷⁷ New South Wales government (2024a).

¹⁷⁸ New South Wales government (2024a).

i Hongkong, genom att de lärarsatta betygen på skolnivå justeras till att motsvara skolans genomsnittliga resultat på HSC-provet (i respektive ämne). Även spridningen kalibreras efter motsvarande spridning i HSC-provet. Elever med de högsta lokala, lärarsatta betygen modereras så att de i princip ges resultatet på skolans högsta examensprovsresultat som modererat betyg (kan alltså vara en annan elevs examensprovsresultat). Detta är ett sätt att hantera tak- och golv-effekter så att elever som går på en skola med svagpresterande kamrater inte ska få sina resultat sänkta. Vissa undantag från moderering görs, t.ex. för sjukdom, men generellt sett utgörs HSC-resultatet av en sammanvägning av det modererade betyget och examensprovsbetyget där de väger 50 % vardera. I examensbeviset redovisas examensprovsresultaten, lärarbetyget, det sammanvägda HSC-betyget samt en redovisning av s.k. *performance bands* (inom vilket intervall för fördefinierade kunskaper inom respektive HSC-ämne som elevens kunskaper uppmätts) för varje ämne.

Rankning av elever för ansökan till högre utbildning

Det finns ett gemensamt australienskt system för rangordning av studenter för antagning till universitet, *Australian Tertiary Admission Rank (ATAR)*.¹⁷⁹ ATAR är ett numeriskt mått på en students övergripande akademiska prestationer i HSC i förhållande till andra studenter och är konstruerat för att möjliggöra jämförelse av prestationer för studenter som har olika kombinationer av HSC-ämnen och kurser. ATAR är inte ett betyg utan beräknas enbart med syfte att användas för antagning, antingen ensamt eller tillsammans med andra urvalskriterier. Beräkningen baseras på antagandet att delstaternas ålderskohorter är jämförbara och ger varje individ ett mått på skalan 0–99,95, där måttet representerar studentens position i dennes ålderskohort. Andelen studerande som antas baserat på ATAR varierar och eftersom antagningen inte är nationellt samordnad har den kritiserats för att vara icke-transparent och försvåra välinformerade val.¹⁸⁰

¹⁷⁹ Universities Admissions Centre (2020).

¹⁸⁰ Higher Education Standards Panel (2016).

International Baccalaureate Diploma Program¹⁸¹

Ett internationellt system för gymnasial utbildning och bedömning i 150 länder

International Baccalaureate Diploma Program är en internationell, tvåårig utbildning som ger examen på gymnasial nivå. IB är en idéburen organisation med ett franchisesystem där skolor som vill erbjuda *Diploma Program* (DP) betalar för rätten och tillgång till material och service från huvudkontoret i Schweiz. I världen finns cirka 5 000 skolor, varav drygt 3 300 på gymnasial nivå, spridda över 150 länder, som erbjuder IB-utbildning. I Sverige har IB-utbildning funnits sedan 1981 och 28 svenska skolor erbjuder i dag DP. En examen från DP ger behörighet till många länders högre utbildning, däribland svenska lärosäten.

En examen kan ge maximalt 45 meritpoäng och beräknas utifrån sex ämnen som betygssätts med betygen 1–7 där 7 är det högsta betyget. Godkäntrörelsen är 24. Kurserna ges i ämnena *studies in language and literature* (språk och litteratur), *language acquisition* (språkinläring), *individuals and societies* (individer och samhälle), *sciences* (naturvetenskap), *mathematics* (matematik) samt *arts* (konsterna). Varje ämne ges på två nivåer *standard level* och *high level*. Även om nivåerna är olika så är betygskriterierna desamma, men eleverna förväntas visa olika djup i sina kunskaper. Därutöver ges tre poäng för *DP core* (kärnelement i DP). Det består av *theory of knowledge (TOK)*, *creativity, activity, service (CAS)* och *extended essay (EE)*, dvs. kunskapsteori, praktiskt utåtriktat arbete och en uppsats.

I *Assessment Principles and Practices* ges en överblick rörande hur bedömning genomförs inom IB, exempelvis hur IB förhåller sig till betygssystemets grund, validitets- och reliabilitetsfrågor, formativ och summativ bedömning, autentiska bedömningssituationer, *backwash*-effekter (hur bedömningen påverkar undervisningen). Avseende *backwash*-effekter lyfts vikten av att skapa bedömningar som stödjer god undervisning och att bedömningarna fungerar för att förutsäga framtida studieframgång.¹⁸²

Grunden för betygssystemet i IB är inte renodlat målrelaterad eller normrelaterad, utan kombinerar dessa två system i något som de kallar för *weak criterion referencing*, dvs. en kombination av mål-

¹⁸¹ Huvudsakliga källor för avsnittet: International Baccalaureate Organization (2018, 2024).

¹⁸² International Baccalaureate Organization (2018, s. 39–45, 72–77 och 90–95).

och normrelativ betygsfördelning. Expertbedömare sätter upp ett spann över tänkta betygsgränser utifrån betygsgränser. Därefter beräknas och justeras betygsgränserna för att överensstämja med tidigare års prestationer så att betygsgränserna länkas mellan åren. Betygssättningen handlar om hur *marks* (poäng) omvandlas till *grades* (betyg). Omvandlingen innebär att betygsgränserna bestäms utifrån en viktning av expertbedömning av betygsunderlaget och kohortens resultat. Det blir således en typ av kravgränssättning när slutbetygen ska sättas i de olika kurserna. Syftet med detta är att kalibrera svårighetsnivån för ett visst betyg. I botten finns en tänkt fördelning för de olika betygen. Viss variation tillåts i fördelningen med motiveringen att IB-systemet är förhållandevis litet och under både tillväxt och utveckling.

Examensprov och slutbetyg

Examensproven utgör en viktig del av IB:s betygssystem. Proven ges vanligen två gånger per år och prövas inte ut på förhand på grund av risken för spridning. Stabiliteten över tid bygger på professionellt omdöme och erfarenhet från provkonstruktörernas sida, men också att kravgränssättningen justeras i efterhand så att en jämn nivå kan bibehållas över tid. Målsättningen är alltså att ett visst provbetyg såväl som ämnesbetyg ska motsvara samma kunskaper i dag som vid vilken annan tidpunkt som helst. Just nu testas genomförande av digitala prov i utvalda skolor. År 2029 väntas IB ha övergått till ett system där samtliga examensprov är digitala.

För elever som av giltig anledning missar ett delprov kan betyg ges baserat på övriga prov och bedömning av exempel på elevens arbete under kursens gång. Det finns också en mekanism för betygssättning när IB inte kan få tillgång till eller bedöma en elevs prestationer. Den används restriktivt när avsaknad av betygsunderlag beror på IB eller tredje part och det inte anses rimligt att eleven ska behöva göra om proven. Ett medelbetyg beräknas då baserat på alla andra ämnen som eleven har fått betyg i de senaste 18 månaderna.

Till skillnad från renodlade examensprovsmodeller är IBO:s hållning att fördelarna överväger nackdelarna med att ämnesbetyg innehåller ett visst mått av lärarbedömning:

There are different views of internal assessment around the world. Some systems place great value on the fact that teachers are best placed to give a holistic opinion on the performance of the candidate rather than just having a “snapshot” from a single assessment. Other education systems rate teacher performance based on their candidates’ outcomes, which creates a strong incentive for teachers to award high marks in any internal assessment [...] The IB believes that the benefits from being able to set meaningful assessment tasks which can only be assessed through classroom-based work which is evaluated by someone who has seen the candidates’ development are greater than the risks internal assessment creates. Therefore, we would usually make an internal assessment task part of any set of assessments.¹⁸³

Lärarbedömning sker i moment som är svårbedömda i skriftliga prov såsom laborationer, muntliga presentationer och grupparbeten. Lärarnas bedömningar granskas sen externt på grund av risken för att olika skolor tolkar betygskriterierna olika och därmed bedömer prestationer olikvärdigt. IB utgår från att lärarnas tolkning av betygskriterierna inom skolan är likvärdig, men för att säkerställa likvärdigheten globalt tillämpas en statistisk modell för att justera betygen mellan skolor. Modereringen utgår från att IB samlar in stickprov av olika lärarbetygsatta underlag som jämförs med huvudbedömarens, utifrån vilka de lärarsatta betygen sedan justeras. Lärarbedömda, modererade moment förekommer som delkomponenter tillsammans med skriftliga prov och viktas olika i olika ämnen. I ämnet *arts* förekommer dock inga skriftliga prov utan endast komponenter som lärarbedöms och modereras. Det kan handla om portföljer med exempel på kreativa elevarbeten, undersökande projekt, skriftliga reflektioner och filmade inspelningar av elevers presentationer. I kärnelementen, EE och TOK, ges betyg A–E, medan CAS inte betygssätts alls. Uppsatsen, EE, handleds på elevens skola men bedöms och betygssätts centralt.¹⁸⁴

Elever kan överklaga sina betyg till sin skola, men skolor kan inte ändra de modererade betygen på eget bevåg. Skolor kan däremot begära att IBO bedömer om hela skolans, men inte enskilda elevers, inskickade material om de misstänker systematiska fel i hanteringen av skolans samlade underlag.

¹⁸³ International Baccalaureate Organization (2018, s. 96).

¹⁸⁴ International Baccalaureate Organization (2016).

Anpassningar vid provgenomförande och regelverk för dessa

Vår genomgång på området omfattar 18 prov i 13 länder och resultaten redovisas detaljerat i översiktstabellerna 10–11. Fem av proven görs i grundskolan och 13 genomförs på gymnasial nivå. Anpassningar vid provgenomförande är möjligt i samtliga länder och alla erbjuder anpassningar vid hörsel- och synnedläggningar. De vanligaste exemplen är anpassningar i form av blindskrift, anpassade utskrifter, undantag från eller särskilt hjälpmedel vid hörförståelseprovdelar eller muntliga prov.

Vissa länder begränsar ansökan om anpassningar vid prov till elever som har anpassningar i sin vanliga skolgång, andra länder tillåter alla att söka. I vissa länder kan man ansöka om specifika anpassningar från en lista medan det i andra länder är upp till den sökande/läraren/läkaren att fritt ange vilka anpassningar som eleven anses behöva. Variationen i ansökningskrav är stor. Ytterligheterna är Hongkong (gymnasial nivå) där ansökan ska göras två år i förväg och Alberta (gymnasial nivå) där ingen ansökan behövs för anpassningar som eleven har i sitt ordinarie skolarbete samt Sverige där formell ansökan helt saknas.

I samtliga fall utom Sverige och Estlands grundskola är det praxis med särskild ansökan i god tid före provgenomförandet. I samtliga fall utom Sverige, Danmark, Norge, Alberta och Estland är det praxis att det krävs ett externt utlåtande från läkare/psykolog eller liknande expertis.

Det vanligaste är att ansökan ska göras till en extern myndighet eller nämnd som beslutar. I Danmark, Norge, Nederländerna, Estland och Alberta kan rektor dock besluta i alla/vissa fall. Möjlighet till undantag från hela prov är mycket ovanligt. Vissa länder nämner explicit möjlighet till undantag från provdelar bl.a. för elever med begränsade språkkunskaper medan det i andra länder inte ens framgår att undantag skulle vara en möjlighet.

Anpassningar inom IB

Den underliggande principen är att anpassningar vid provgenomförande inom IB ska medges för elever som har anpassningar även i den ordinarie undervisningen.¹⁸⁵ Sjukdom och tillstånd som varar mer än 12 veckor kan föranleda bedömning av behov av anpassningar för att underlätta en elevs lärande. För oförutsedda tillstånd av annan karaktär, både händelser som drabbar enskilda och naturkatastrofer eller liknande som drabbar grupper, har IBO ett annat regelverk än det för anpassningar.

Beslut om anpassningar tas centralt av IB, ska vara strikt individinriktat och elevens behov ska följas upp över tid. Det innebär att beslut kan överklagas och ändras åt båda håll.

Vilka anpassningar som kan övervägas vid olika tillstånd är tydligt specificerat inklusive vilka som kan vara aktuella i vissa sammanhang och i vilken omfattning. Det kan både i undervisnings- och provsituationer röra sig om exempelvis hjälpmedel för syn- och hörselnedsättning, extra tid, raster och hjälp av en *scribe*, dvs. en assistent som skriver ner ens dikterade svar. Senarelagd examination kan också medges. Muntlig framställan, såsom att få spela in sina svar på uppgifter eller prov i stället för att skriva ner dem medges inte. Undantag från genomförande av enskilda moment kan tillåtas i exceptionella fall såsom hörförståelse för personer med hörselnedsättning.

Ansökan om anpassningar ska innehålla en medicinsk och/eller psykologisk bedömning, ett utlåtande från skolan samt resultatet på ett språktest för elever som läser utbildningen på ett annat språk än sitt modersmål. Beslut ska baseras på lärarens observationer av eleven i klassrummet och en viss anpassning kan inte tillhandahållas som standard till alla elever med behov av stöd för lärande i skolan/klassrummet.

Andelarna underkända och obehöriga elever i andra länder

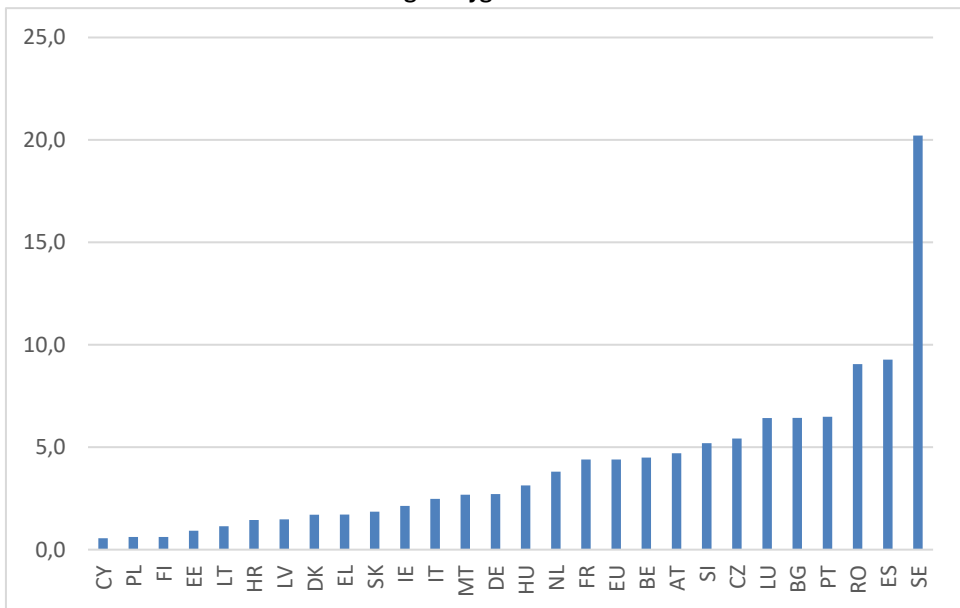
Av genomgången av enskilda länder framgår att majoriteten av länder inte har någon formell skarp godkäntgräns i de lärarsatta betygen eller behörighetskrav till utbildning på gymnasial nivå. I detta avsnitt

¹⁸⁵ International Baccalaureate Organization (2022).

jämför vi hur stor andel av elever som inte blir behöriga till nästa utbildningsnivå i Sverige jämfört med andra länder i relation till prestationsnivå i de internationella studierna.

Vid jämförelse med andra länder avseende andelen underkända betyg och andelen elever som inte blir behöriga till fortsatt utbildning framgår att Sverige ligger högt (figur 1). Genomsnittet för EU ligger strax under fem procent, medan Sverige ligger på drygt 20 procent.¹⁸⁶ Det land som kommer näst efter Sverige är Spanien på knappt tio procent. Det innebär att Sverige 2018 låg nästan fyra gånger högre än genomsnittet i EU och mer än dubbelt så högt som det land som låg näst högst vad gäller andel obehöriga till gymnasial utbildning.

Figur 1 Andel elever som går ut den obligatoriska skolans sista årskurs utan fullständiga betyg



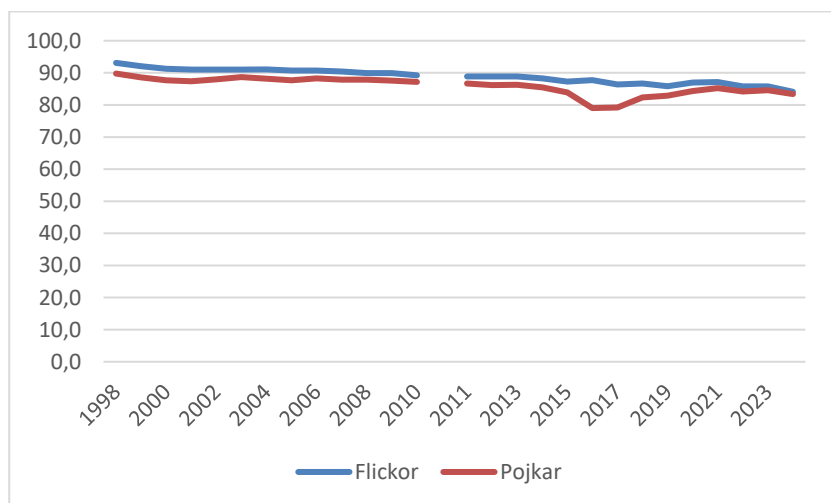
Källa: OECD och Pisa 2018.

Anm. Resultaten bygger på uppgifter från rektorer. Sveriges stapel betecknar andelen elever i grundskolans sista årskurs som gick ut skolan utan behörighet till nationella program i gymnasieskolan.

¹⁸⁶ Figuren bygger på data insamlad genom en urvalsundersökning genomförd av PISA så den kan avvika från den offentliga statistiken. Även enligt den offentliga statistiken ligger Sverige betydligt högre än jämförbara länder.

Figur 1 bygger på data som kommer från PISA 2018 och det är enligt Skolverkets statistik inte fullt så många som blev obehöriga till ett nationellt program i gymnasieskolan runt de åren (se figur 2). Även med de officiella uppgifterna från Skolverket över andelen obehöriga i riket som helhet så har Sverige en hög andel elever som ej blir behöriga eller har fullständiga betyg från grundskolan jämfört med andra EU-länder.

Figur 2 Andel (procent) elever som är behöriga till gymnasieskolans nationella program, flickor och pojkar, vårterminerna 1998–2024



Källa: Skolverket (2023c).

Den höga andelen F-betyg i Sverige kan ställas i relation till de internationella studierna TIMSS 2019 för matematikämnet respektive PISA 2022 för områdena matematik, läsförståelse och naturkunskap där Sverige presterade bättre än OECD-genomsnittet.

I PISA presterade 27 procent av eleverna i Sverige på nivå 2 eller lägre i matematik, OECD-genomsnittet låg på 31 procent. I läsförståelse låg Sverige på 24 procent som presterade på nivå 2 eller lägre, OECD-genomsnittet var på 26 procent. I naturvetenskap låg Sverige på 24 procent som presterade på nivå 2 eller lägre, vilket var detsamma som OECD-genomsnittet.¹⁸⁷

¹⁸⁷ Skolverket (2023a).

I TIMSS 2019 presterade Sverige under OECD-genomsnittet på totalpoängen i matematik.¹⁸⁸ Det kan dock vara rimligt att jämföra andelen lågpresterande i matematik som var 10 % i TIMSS 2019 med andelen obehöriga i figur 1, eftersom matematik är det ämne som är vanligast att svenska elever inte har godkänt betyg i förutom svenska som andraspråk (dock med ett betydligt lägre antal elever som läser ämnet).¹⁸⁹ Sveriges andel lågpresterande ligger på samma nivå som EU- och OECD-genomsnittet i TIMSS 2019 och lägre än t.ex. Cypern som har den lägsta andelen obehöriga i EU. Finland hade marginellt lägre andel lågpresterande i TIMSS 2019, men ligger på väldigt låg andel elever som inte går vidare i utbildningssystemet. Frankrike har en större andel lågpresterande än Sverige men ligger också mycket lägre avseende andelen som inte går vidare.

Sveriges resultat i ICCS som mäter elevers kunskaper, attityder och engagemang i medborgar-, demokrati- och samhällsfrågor i årskurs 8 är bland de bästa resultaten av de deltagande länderna.¹⁹⁰ Även i SO-ämnena har Sverige dock en relativt hög andel F-betyg.

Andel elever som inte klarar gymnasieskolan och gymnasieexamen

Bland de länder som vi har inkluderat i jämförelsen har Sverige en relativt hög andel elever som påbörjar gymnasieskolan, men en relativt låg andel som tar en gymnasieexamen, se figur 3.

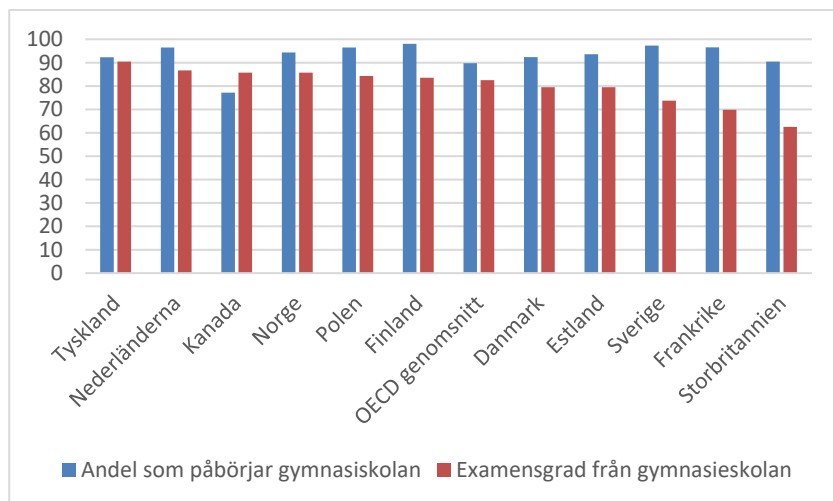
¹⁸⁸ Skolverket (2020).

¹⁸⁹ Skolverket (2023a).

¹⁹⁰ Skolverket (2023b).

Figur 3 Andel elever som påbörjar gymnasieskolan resp. examensgrad från gymnasieskolan år 2020, %

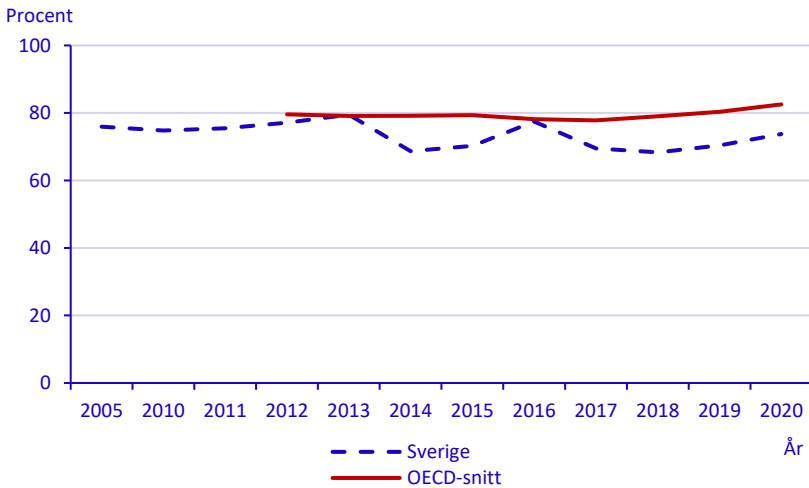
Graden av gymnasieexamen representerar den uppskattade andelen personer som kommer att ta examen från gymnasieutbildningen under sin livstid



Källa: OECD.

Det framgår av föregående avsnitt att Sverige utmärker sig med genomsnittliga resultat i de internationella studierna, men en mycket hög andel underkända/elever som inte blir behöriga till gymnasieskolan eller motsvarande. Även avseende andel elever som tar en gymnasieexamen, dvs. klarar en gymnasieutbildning med de resultat som krävs för examen, tycks Sverige ha relativt högt ställda krav.

Figur 4 Genomströmning i gymnasieskolan
Avklarandegrad av gymnasieexamen 2005–2020 i Sverige
och OECD-genomsnittet



Källa: SCB (2024).

I figur 4 jämförs trenden i Sverige med OECD angående avklarad gymnasieexamen. I Sverige kan vi se nedgångar i examensgraden efter införandet av Gymnasiereformerna 2011 med hårdare krav för examen, och efter flyktingströmmen 2015–2016.

Översiktstabeller

I detta avsnitt presenteras översiktstabeller över undersökta aspekter för länderna i jämförelsen.

Tabell 4 Bakgrundsvariabler

	Offentliga utgifter för utbildning i grundskola och på gymnasial nivå, % av BNP	Utbildningsstruktur*	Undervisning per år i grundskolans senare årskurser**, antal timmar	15-åriga elever som rapporterar känsla av att höra hemma i skolan, pojkar/flickor %	OECD-rankning social jämlikhet i utbildning, placering av 39 länder	Elever 13–15 år som uppger erfarenhet av mobbning senaste månaden, pojkar/flickor %	PISA-rankning 2022 (2018)
Sverige	4,1	SS	869	70/64	16	18/19	19 (17)
Finland	3,8	SS	808	77/74	9	27/24	13 (10)
Danmark	3,7	SS	1 200	73/71	12	12/15	17 (18)
Norge	4,7	SS	874	78/76	8	16/14	32 (23)
Nederländerna	3,1	TS	1 000	73/79	19	12/12	25 (16)
Frankrike	3,5	PS	968	37/39	32	13/16	26 (26)
Polen	2,9	PS	754	57/62	17	23/18	12 (11)
Estland	3,2	SS	823	73/74	1	NA	7 (5)
England	3,7	PS	Ej reglerat	65/60	11	38/37	20 (13)
Tyskland	2,9	TS	896	74/76	34	21/21	24 (18)
Hongkong	–	PS	–	–	NA	–	6 (4)
Singapore	–	PS	–	–	NA	–	1 (2)

	Offentliga utgifter för utbildning i grundskola och på gymnasial nivå, % av BNP	Utbildningsstruktur*	Undervisning per år i grundskolans senare årskurser**, antal timmar	15-åriga elever som rapporterar känsla av att höra hemma i skolan, pojkar/flickor %	OECD-rankning social jämlikhet i utbildning, placering av 39 länder	Elever 13–15 år som uppger erfarenhet av mobbning senaste månaden, pojkar/flickor %	PISA-rankning 2022 (2018)
Kanada	3,4	PS	925	70/65	2	36/40	8 (8)
Australien	3,7	PS	1 000	70/66	18	–	11 (21)
OECD	3,3	–	916	71	–	26/26***	–

Källor: OECD (2023a); OECD (2024a); OECD (u.å.a); OECD (u.å.b); UNICEF (2023).

Anm.: * SS = Single Structure, en sammanhållen grundskola; PS = Primary/Secondary, grundskolans högre årskurser tillhör sekundärskolan ihop med gymnasieskolan; TS = Tracked Secondary, uppdelning i inriktningar redan till sekundärskolan baserat på kunskaper/betyg, lärarbedömning eller intressen.

** OECD:s definition och klassificering av *lower secondary school*.

*** Europa och Centralasien.

Tabell 5 Utmärkande drag för grundskolan

	Obligatorisk förskola?	Ålder för skolstart	Antal år obligatorisk utbildning	Privatskolesystem?
Sverige	NEJ	6	10	16 % elever. Ingen avgift, får ta ut vinst.
Danmark	JA	7	10	18 % elever. Avgift, får ej ta ut vinst.
Norge	NEJ	6	10	3 % elever. Avgift, får ej ta ut vinst.
Finland	JA	7	11	2 % elever. Ingen avgift, får ej ta ut vinst.
Estland	NEJ	7	9	4 % elever. Avgift varierar, får ta ut vinst.
Nederländerna	JA	6	11	70 % elever. Avgift förekommer, får ej ta ut vinst.
Frankrike	JA	6	12	20 % elever. Avgift förekommer, får ej ta ut vinst.
New South Wales	NEJ	6	11	35 % elever. Avgift, får ta ut vinst med begränsningar, beroende på huvudman.
Polen	JA	7	9	7 % elever. Avgift, får ta ut vinst.
Alberta	NEJ	6	12	6 % elever. Avgift, får ta ut vinst.
Hongkong	NEJ	6	9	Låg andel. Främst internationella elever. Avgift, får ta ut vinst.
Singapore	NEJ	6	10	Låg andel. Främst internationella elever Avgift, får ta ut vinst.
England	NEJ	5	11/13	6 % elever. Avgift, får ej ta ut vinst.
Nordrhein-Westfalen	NEJ	6	12/13	9 % elever. Avgift förekommer, får ej ta ut vinst.

Källor: Se textavsnitt för respektive land.

Tabell 6 Betyg i grundskolan

	Slutbetyg i grundskolan	Betygsskala (lägst till högst)	Godkänthögräns i betygs- skalan	Går betyg att överklaga?	Betyg i ordning och uppförande?
Sverige	Lärarsatta betyg	F–A	E	NEJ	NEJ
Danmark	Sammanvägning lärarsatta betyg och slutprov	F–A inkl. Fx	E	JA	Frivilliga försök pågår
Norge	Lärarsatta betyg och slutprov	1–6	Saknas	JA	JA*
Finland	Lärarsatta betyg	1–10	5	JA	JA**
Estland	Lärarsatta betyg och slutprov	1–5	3	JA	NEJ
Nederländerna	Sammanvägning lärarsatta betyg 50 %, slutprov 50 %	1–10	6	NEJ	NEJ
Frankrike	Sammanvägning, lärarsatta betyg 50 %, examensprov 50 %	1–20	Ej skarpt men oftast 10	NEJ	JA
New South Wales	Lärarsatta betyg	E–A	E	NEJ	NEJ
Polen	Lärarsatta betyg och examensprov	1–6	2	JA	JA
Alberta	Lärarsatta betyg	Varierar	Varierar	Varierar	NEJ
Hongkong	Lärarsatta betyg	F–5**	2–5**	Framgår inte	Framgår inte
Singapore	Examensprov	0–100	Varierar	JA	JA
England	Examensprov	1–9	Varierar, oftast 4/5	JA	NEJ
Nordrhein-Westfalen	Lärarsatta betyg	6–1	4	JA	NEJ

Källor: Se textavsnitt för respektive land.

Anm.: * Räknas ej vid urval till fortsatt utbildning men används vid lärlingsansökan. ** Anges ej i slutbetyget.

Tabell 7 Centrala prov i grundskolan

	Centrala prov	Syfte med proven	Prov-konstruktion	Skriftliga prov	Muntliga prov	Digitala prov	Rättning	Provbetygskala (lägst till högst)
Sverige	JA	Stöd för betygs-sättning i fem ämnen	Lärosäten	JA	JA	NEJ	Lärare	F–A
Danmark	JA	Formativa, diagnostiska och examens-prov	Centrala kommittéer	JA	JA	JA	Automatiserad för kvalitetssäkring. Extern för examens-prov.	F–A inkl. Fx
Norge	JA	Examensprov	Lokalt	JA	JA	JA	Extern	1–6
Finland*	NEJ	–	–	–	–	–	–	–
Estland	JA	Examensprov	Utbildnings-ministeriet	JA	JA	JA	Automatiserad, central och lärarbaserad (anonymiserad)	1–100 %
Nederländerna	JA	Examensprov	Central, Cito	JA	JA	NEJ	Två lärare, en från den egna skolan och en från en annan skola.	1–10
Frankrike	JA	Examensprov	Framgår ej	JA	JA	NEJ	Extern, andra lärare	0–400
New South Wales	JA	Diagnostiskt	Central	JA	NEJ	JA	Automatiserad och central	Poäng som räknas om till fyra nivåer
Polen	JA	Examensprov	Central	JA	NEJ	NEJ	Extern	Framgår ej

	Centrala prov	Syfte med proven	Prov-konstruktion	Skriftliga prov	Muntliga prov	Digitala prov	Rättning	Provbetygskala (lägst till högst)
Alberta	JA	Formativa/ diagnostiska	Central	JA	NEJ	JA	Lärare och centralt	1–5/uppgift som sen räknas om till 1–100 %
Hongkong	JA	Formativa/ diagnostiska	Central	JA	JA	NEJ	Extern	Ingen återkoppling till individen
Singapore	JA	Formativa, diagnostiska och examens- prov	Central	JA	JA	JA	Extern	0–100 som omvandlas till 1–8
England	JA	Examensprov	Fem olika <i>exam boards</i>	JA	JA	NEJ	Extern	1–9
Nordrhein- Westfalen	JA	Formativa och diagnostiska "learning status surveys"	Central, IQB	JA	JA	Varierar mellan delstater	Lärare	I–V där III ska motsvara förväntad nivå

Källor: Se textavsnitt för respektive land.

Anm.: * Urvalsbaserade prov för nationell trendmätning förekommer men återkoppling sker inte på individnivå.

Tabell 8 **Gymnasieskola och betyg**

	Godkäntgräns för tillträde till gymnasiestudier?	Slutbetyg i högskoleförberedande gymnasieskola	Betygsskala slutbetyg (lägst till högst)	Går betyg att överklaga?	Betygs-genomsnitt huvudgrund för urval till högskola?	Kan yrkesgymnasium kvalificera för högskola (ej yrkeshögskola)?
Sverige	JA	Lärarsatta	F–A	NEJ	JA	JA
Danmark	JA	Lärarsatta + slutprovsresultat	F–A	JA	JA	Förekommer
Norge	NEJ	Lärarsatta + slutprovsresultat	1–6 (1 = underkänt)	JA	JA	NEJ, kräver påbyggnadsår
Finland	JA, men möjlighet till prövning finns också	Studentexamen/examensprov	4–10 men vissa kurser IG/G (4 = underkänt)	JA	Varierar	Förekommer
Estland	NEJ	Lärarsatta + slutprovsresultat	2–5 men vissa kurser IG/G	JA	Varierar	Förekommer, men kan kräva påbyggnad
Nederländerna	NEJ	Slutprov	1–10	NEJ	NEJ	NEJ
Frankrike	NEJ	Lärarsatta 40%, examensprov 60%	1–20	NEJ	NEJ	NEJ
New South Wales	NEJ	Provresultat, lärarsatta betyg som modereras utifrån provresultat, HSC-mark samt <i>performance band</i>	E–A /ATAR-poäng	JA	JA	Förekommer, men kan kräva påbyggnad
Polen	NEJ	Lärarsatta + slutprovsresultat	1–6	JA	NEJ	Kräver längre kombinationsvariant

	Godkänträns för tillträde till gymnasiestudier?	Slutbetyg i högskoleförberedande gymnasieskola	Betygsskala slutbetyg (lägst till högst)	Går betyg att överklaga?	Betygs-genomsnitt huvudgrund för urval till högskola?	Kan yrkesgymnasium kvalificera för högskola (ej yrkeshögskola)?
Alberta	Framgår ej	Lärare 70%, examensprov 30%	0–100 %	JA	Varierar	Framgår ej
Hongkong	Varierar	HKDSE-examensprov	2–5**	NEJ	Varierar	Framgår ej
Singapore	Varierar	JA, olika varianter av examensprov	Varierar	JA	JA	Framgår ej
England	NEJ	A-levels, examensprov	E–A*/UCAS-poäng	JA	JA	NEJ
Nordrhein-Westfalen	Inte formellt, men lärarekommendation väger tungt från åk 5	Abiturbetyg + lärarbetyg	7-gradig men med steg 0–15 lärarsatta betyg i vissa delar	JA	JA	Förekommer

Källor: Se textavsnitt för respektive land.

Anm.: Notera att asteriskerna i Hongkongs och England betygsskalor betecknar betygssteg. I Hongkong finns, utöver betygssteget 5, stegen *Five-Star* och *Five-Double-Star*. I England infördes betygssteget A* (*A-star*) år 2010 för att bättre kunna differentiera mellan högpresterande elever vid ansökan till högskolan.

Tabell 9 *High stakes-prov i gymnasieskolan*

	Syfte	Muntliga prov	Digitala prov	Rättning	Provbetygskala (lägst till högst)
Sverige	Stöd för lärares betygs-sättning i tre ämnen	JA	NEJ	Lärare	F–A
Danmark	Examensprov	JA	JA	Extern	F–A
Norge	Examensprov	JA	JA	Extern	1–6 (1=underkänt)
Finland	Examensprov	NEJ	JA	Lärare + extern	7 steg
Estland	Examensprov	I språk	JA	Automatiserad, central och lärarbaserad (anonymiserad)	1–100 %
Nederländerna	Examensprov	JA	JA	Automatiserad, lärare + extern	1–10
Frankrike	Examensprov	JA	NEJ	Extern	1–20
New South Wales	Examensprov	JA	Delvis	Extern	0–100
Polen	Examensprov	JA	NEJ	Extern	0–100
Alberta	Examensprov	NEJ	JA	Extern	0–100
Hongkong	Examensprov	JA	NEJ	Extern	1–5**
Singapore	Examensprov	JA	Delvis	Extern	Olika
England	Examensprov	JA	NEJ	Extern	E–A*
Nordrhein-Westfalen	Examensprov	JA	NEJ	Två lärare, en från egen skola och en från en annan skola.	1–300

Källor: Se textavsnitt för respektive land.

Tabell 10 Anpassningar vid prov

	Tillstånd som medger anpassningar vid prov	Vanligaste exemplen på anpassning	Krävs formell ansökan/blankett?	Krävs externt utlåtande?	Vem beslutar?
Sverige, grundskola och gymnasieskola	Funktionsnedsättning, läs- och skrivsvårigheter. Bristande kunskaper i svenska.	Förlängd provtid, dela upp provet. Provet ska fortfarande pröva det som är syftet och lärarens bedömning får inte anpassas.	NEJ	NEJ	Rektor
Danmark, grundskola	Fysiskt/psykiskt funktions- hinder. Bristande kunskaper i danska.	Tekniska hjälpmedel, förlängd provtid, en assistent. Svårighetsgraden får inte minska.	JA	Kan krävas, men rektor avgör om det behövs.	Rektor
Estland, grundskola	Elever som på grund av behov av specialpedagogiskt stöd har beslut om anpassningar/hjälp- medel i vanliga fall.	Förlängd provtid, pauser, eget rum, assistent, individuell anpassning p.g.a. begränsad språkkunskap eller individuell studieplan beslutad sedan tidigare.	För vissa anpassningar krävs ansökan, andra kan rektor besluta om.	Nej	Rektor resp. skol- myndigheten Harno.
Polen, grundskola	Elever med särskilda utbild- ningsbehov, inklusive elever med funktionshinder, social missanpassning och risk för social missanpassning.	Tillåtna anpassningar meddelas årligen av utbildningsstyrelsen, i ett utlåtande från direktören för CKE, Centralna Komisja Egzaminacyjna (Central Examination Commission). Utlåtandet publiceras senast den 20 augusti det läsåret som föregår det läsåret under vilket det är examen i åttonde klass.	JA, till district examination board för mycket speciell typ av anpassning.	JA, senast 15 oktober.	Rektor/ <i>exam board</i> beslutar på underlag av att huvudläraren föreslår föräldrarna vilka anpassningar som kan göras. Rektor meddelar senast 20 nov. vad <i>exam board</i> föreslår.

	Tillstånd som medger anpassningar vid prov	Vanligaste exemplen på anpassning	Krävs formell ansökan/blankett?	Krävs externt utlåtande?	Vem beslutar?
Frankrike, grundskola och gymnasieskola	Begränsning av aktiviteter eller deltagande i samhällslivet på grund av en betydande, långvarig eller permanent nedsättning av en eller flera fysiska, sensoriska, mentala, kognitiva eller psykiska funktioner, ett flerk Funktionshinder eller en funktionsnedsättande hälsostörning.	Förlängd provtid, pauser, eget rum, assistent, sprida ut proven över tid.	JA, inkl. lärares och föräldrars info. Enklare blankett för elever som redan har någon form av anpassning beviljad.	JA, läkare	Examensarrangören
England, grundskola och gymnasieskola	Fysisk/psykiskt Funktionshinder.	Extra provtid, undantag från provdel, uppläsning, digitalt prov i stället för på papper och tvärtom. IT-hjälpmiddel. Elever med samma diagnos får inte alltid samma anpassning utan det görs en individuell bedömning. Tillfällig sjukdom/olycka mm behandlas i separat ordning.	JA, skolan söker till den Exam Board som eleven ska genomföra prov hos.	JA	Exam Board
Danmark, gymnasieskola	Fysiskt/psykiskt Funktionshinder.	Förlängd provtid, paus, diktering, assistent. Svårighetsgraden får inte minskas. Som tumregel kan anpassningar som en elev får använda i sin vanliga undervisning användas vid proven så länge svårighetsgraden inte minskas.	JA	Skolledningen avgör	Skolledningen

	Tillstånd som medger anpassningar vid prov	Vanligaste exemplen på anpassning	Krävs formell ansökan/blankett?	Krävs externt utlåtande?	Vem beslutar?
Norge, gymnasieskola	Upp till rektor att avgöra. Även elever som inte har haft behov (vedtak om specialundervisning) tidigare kan ansöka.	Förlängd provtid, paus, diktering, assistent, särskilt IT-hjälpmedel.	JA	Rektor avgör	Rektor
Finland, gymnasieskola	Sjukdom och funktionshinder, Läs- och skrivsvårigheter, Bristfälliga färdigheter i undervisningsspråket hos examinander med främmande språk som modersmål, Synnerligen svår livssituation.	Extra provtid, en enskild provlokal, större teckenstorlek, större bildskärm, prov som avläggs med hjälpmedel för synskadade, prov med begränsat ljudmaterial, utskrift i stället för digitalt prov eller anlåtande av assistent. Även generösare bedömning kan tillämpas vid rättning.	JA, rektor söker åt elev.	JA, både externt och av undervisande lärare	Studentexamensnämnden
Estland, gymnasieskola	Elever som på grund av behov av specialpedagogiskt stöd har beslut om anpassningar/hjälpmedel i vanliga fall.	Förlängd provtid, pauser, eget rum, assistent, individuell anpassning till följd av begränsad språkkunskap eller individuell studieplan beslutad sedan tidigare.	JA, skolan ansöker till exam board	Framgår inte	Skolmyndigheten Harno
Nederländerna, gymnasieskola	Funktionshinder inkl. NPF. Färre än 6 års skolgång i Nederländerna.	30 min förlängd provtid. Uppläsning av text. Ordbok vid annat modersmål.	JA	JA, psykolog, läkare, neurolog	Skolan, som dock kan be skolmyndigheten CVTE om råd.

	Tillstånd som medger anpassningar vid prov	Vanligaste exemplen på anpassning	Krävs formell ansökan/blankett?	Krävs externt utlåtande?	Vem beslutar?
New South Wales, gymnasieskola	Funktionshinder, i generell och bred bemärkelse inkl. tillfällig skada/sjukdomstillstånd.	Assistent, syn/hörtolk, raster, diabeteshjälpmiddel.	JA, inkl. att rektor och skolpersonal måste fylla i delar.	JA. Om det saknas behöver det beskrivas varför eleven inte kan få en diagnos.	Skolmyndigheten NESA uppges behandla 10 000 ansökningar per år.
Polen, gymnasieskola	Fysiska och mentala funktionshinder, bristande språkfärdigheter, inlärningssvårigheter, tillfällig skada/sjukdom, kris, trauma, språkstörning.	Förlängd provtid, särskilda provformulär, sitta avskilt, assistent. Vid hjälp av assistent ska ljudinspelning ske för att kunna kontrollera för otillåten hjälp.	JA. Lärare (som måste ha föräldrars tillåtelse) ber rektor som ansöker till Examination Board.	JA, extern psykolog/specialpedagog	Regional Examination Board
Alberta, gymnasieskola	Utgår från vad eleven har beviljats för hjälpmedel i vanliga fall.	Extra tid, pauser, uppläsning, assistent (scribe/reader).	NEJ, inte för anpassningar som eleven har i vanliga fall. JA, för alla övriga anpassningar, rektor söker minst 2 månader i förväg.	NEJ	Rektor/provinsen
Hongkong, gymnasieskola	Fysiska funktionshinder, NPF, inlärningssvårigheter.	Extra tid, pauser, förstörd text, IT-hjälpmiddel, undantag från enskild provdel. Anpassning får ej medföra fördel jämfört med andra provtagare.	JA, två år i förväg.	JA	Committee on Special Needs Candidates (CSNC) och Committee on Candidates with Specific Learning Disabilities (CSLD) enligt de riktlinjer som fastställts av Public Examinations Board.

	Tillstånd som medger anpassningar vid prov	Vanligaste exemplen på anpassning	Krävs formell ansökan/blankett?	Krävs externt utlåtande?	Vem beslutar?
Singapore, gymnasieskola	Inlärningssvårigheter, fysiskt funktionshinder.	Extra tid, undantag från provdel, ändrad provdel, uppläsning, assistent (scribe/reader). Betyget innehåller markeringar för vilka anpassningar som gjorts för resp. provdel.	JA, skolan söker s.k. AA (access arrangements) till central enhet, SEAB i mycket god tid.	JA, ska bifogas ansökan.	Skolmyndigheten SEAB, centralt. Besked tar 3 månader att få.
Nordrhein-Westfalen, gymnasieskola	Funktionsnedsättning, diagnostiserad kronisk sjukdom eller svårighet/störning, inklusive elever med autism som inte behöver särskilt pedagogiskt stöd i vanliga fall. Även tillfällig skada/sjukdom kan medge anpassning.	Extra provtid, raster, tekniska hjälpmedel, eget rum, assistent. Exempler på anpassningar nämns som generella anpassningar, dvs. inte för just abiturprov. Elever med funktionshinder ska ges stöd i att välja ämnen för sina gymnasiestudier så att de väljer med hänsyn tagen till att de ska kunna genomföra examensprov i ämnena. Dyskalkyli erkänns inte som funktionshinder och medger aldrig anpassningar.	JA. Skolledningen söker till skolmyndigheten Obere Schulaufsicht senast efter höstlovet det år som föregår Abitur-examen.	Framgår inte	Obere Schulaufsicht, skolmyndighet på delstatsnivå

Källor: Se textavsnitt för respektive land.

Tabell 11 Undantag från provgenomförande

	Kan undantag från prov medges?	Tillstånd som medger undantag från prov
Sverige, grundskola och gymnasieskola	JA, från delar eller hela prov.	Särskilda skäl. Bristande kunskaper i svenska.
Danmark, grundskola	JA	Om anpassningar inte bedöms räcka, föräldrar och elev samtycker samt en alternativ bedömningsform tas fram för eleven.
Estland, grundskola	Framgår inte	
Polen, grundskola	JA	En elev som har ett intyg om eller utreds för måttlig eller svår utvecklingsstörning, en kombinerad funktionsnedsättning där en av funktionsnedsättningarna är en måttlig eller svår intellektuell funktionsnedsättning. En elev kan också undantas till följd av särskild tillfällighet eller av hälsoskäl som omöjliggjort genomförande vid både ordinarie och uppföljande provtillfälle.
Frankrike, grundskola och gymnasieskola	JA, från provdelar, men inte hela prov.	Framgår inte,
England, grundskola och gymnasieskola	Framgår inte	–
Danmark, gymnasieskola	NEJ	–
Norge, gymnasieskola	NEJ	–
Finland, gymnasieskola	NEJ	–

	Kan undantag från prov medges?	Tillstånd som medger undantag från prov
Estland, gymnasieskola	Framgår inte	–
Nederländerna, gymnasieskola	Framgår inte	–
New South Wales, gymnasieskola	JA, kan ske helt eller delvis för att ersättas av andra bedömningsformer, men även det beslutas av NESAs.	Framgår inte.
Polen, gymnasieskola	JA	Hörselnedsättning medger undantag från hörförståelse. Stumhet medger undantag från muntliga prov.
Alberta, gymnasieskola	JA, kräver ansökan till provinsens skolmyndighet. Kan ej beviljas av rektor. Kräver föräldrars samtycke.	Sjukdom, dödsfall, väder, samhällelig kris, kallad till rätten.
Hongkong, gymnasieskola	JA, från enskilda provdelar.	Framgår inte.
Singapore, gymnasieskola	JA	Framgår inte helt, men hörselnedsättning nämns.
Nordrhein-Westfalen, gymnasieskola	Framgår inte	–

Källor: Se textavsnitt för respektive land.

Referenser

- ACARA (2023). *National Report on Schooling in Australia. Student numbers*. <https://www.acara.edu.au/reporting/national-report-on-schooling-in-australia/student-numbers> [2024-08-07].
- ACARA (2024). *NAPLAN*. <https://www.nap.edu.au/naplan> [2024-08-07].
- Beau, A. (2022). Taux de réussite, filières, mentions... Les résultats du bac 2022 en trois infographies. *L'Express*. https://www.lexpress.fr/societe/education/taux-de-reussite-filieres-mentions-les-resultats-du-bac-2022-en-trois-infographies_2176797.html [2024-07-08].
- Bekendtgørelse om obligatoriske test i folkeskolen i skoleårene 2022/23 til og med 2025/26 (Folkeskolens Nationale Overgangstest). *Styrelsen for Undervisning og Kvalitet*, den 31. august 2022. <https://www.retsinformation.dk/eli/lta/2022/1223> [2024-07-15].
- Børne- og Undervisningsministeriet (2023a). *Nu kan skoler deltage i forsøg med karakterfrihed eller nye karakterer*. <https://www.uvm.dk/aktuelt/nyheder/uvm/2023/feb/230207nu-kan-skoler-deltage-i-forsog-med-karakterfrihed> [2024-07-15].
- Børne- og Undervisningsministeriet (2023b). *Beviser og beståkrav*. <https://www.uvm.dk/folkeskolen/folkeskolens-proever/bedoemmelse-og-censur/beviser-og-bestaakrav> [2024-07-15].

- Børne- og Undervisningsministeriet (2024). *Folkeskolens Nationale Overgangstest*. <https://www.uvm.dk/folkeskolen/test-evaluering-og-skoleudvikling/test--og-evalueringsredskaber/folkeskolens-nationale-overgangstest/folkeskolens-nationale-overgangstest> [2024-07-12].
- Calmand, J., Giret, J. F., Guégnard, C. & Paul, J. J. (2009). Youth transitions at risk? Insecurity, precarity and educational mismatch in the youth labour market. Why Grandes Écoles are so valued? *17th annual workshop of the European Research Network on Transitions in Youth IREDU, Burgundy University.*, September 2009, Dijon.
- Cambridge University Press & Assessment (2022). *Special Consideration*. <https://www.cambridgeinternational.org/covid/june2022-exam-series/running-exams/special-consideration/> [2024-07-15].
- Cambridge University Press & Assessment (2024). *Calculating assessed marks*. <https://www.cambridgeinternational.org/covid/june2022-exam-series/running-exams/calculating-assessed-marks/> [2024-07-15].
- Centralna Komisja Egzaminacyjna (2024a). *O egzaminie*. [Om proven] <https://cke.gov.pl/egzamin-maturalny/egzamin-maturalny-w-formule-2023/o-egzaminie/> [2024-07-16].
- Centralna Komisja Egzaminacyjna (2024b). *Podstawa programowa* [Läroplan]. <https://cke.gov.pl/egzamin-maturalny/egzamin-maturalny-w-formule-2023/podstawa-programowa/> [2024-07-16].
- Conference des Grandes Écoles (2023). *Accès aux Grandes écoles*. <https://www.cge.asso.fr/commissions/acces-aux-grandes-ecoles/> [2024-07-16].
- Council of Ministers of Education, Canada (u.å) *PCAP*. [https://www.cmec.ca/240/Programs-and-Initiatives/Assessment/Pan-Canadian-Assessment-Program-\(PCAP\)/Overview/index.html](https://www.cmec.ca/240/Programs-and-Initiatives/Assessment/Pan-Canadian-Assessment-Program-(PCAP)/Overview/index.html) [2024-07-17].
- Council of Ministers of Education, Canada (2023). *Secondary Education in the Provinces and Territories of Canada: A Student Transfer Guide 2023–24*.

- https://cmec.ca/docs/transferguide/2024/Alberta-2023_Student-Transfer-Guide_EN.pdf [2024-11-01].
- Creese, B., Gonzalez, A. & Isaacs, T. (2016). Comparing international curriculum systems: The international instructional systems study. *The Curriculum Journal*, 27(1), 523.
<http://dx.doi.org/10.1080/09585176.2015.1128346>.
- Department for Education (2023). *Introduction of T Levels*. Policy Paper. <https://www.gov.uk/government/publications/introduction-of-t-levels/introduction-of-t-levels> [2024-07-15].
- Education Bureau (2023). *Secondary School Places Allocation System 2022/2024. (For Admission to Secondary 1 in September 2024): Frequently Asked Questions*.
https://www.edb.gov.hk/attachment/en/edu-system/primary-secondary/spa-systems/secondary-spa/general-info/FAQ_SSPA_2024_Sept_General_EN.pdf [2024-07-17].
- Education GPS (u.å.). OECD. Alberta (Canada)
<http://gpseducation.oecd.org> [2024-07-17].
- Eurydice (2023a). *France. 3.2 Higher education funding*.
<https://eurydice.eacea.ec.europa.eu/national-education-systems/france/higher-education-funding> [2024-07-16].
- Eurydice (2023b). *Netherlands. 6.3 Assessment in general secondary education (HAVO, VWO)*.
<https://eurydice.eacea.ec.europa.eu/national-education-systems/netherlands/assessment-general-secondary-education-havo-vwo> [2024-07-16].
- Eurydice (2023c). *Netherlands. 7.2.1 Bachelor*.
<https://eurydice.eacea.ec.europa.eu/national-education-systems/netherlands/bachelor> [2024-07-16].
- Eurydice (2024a). *Norway. 6 Upper secondary education and post-secondary tertiary education*.
<https://eurydice.eacea.ec.europa.eu/national-education-systems/norway/upper-secondary-education-and-post-secondary-tertiary-education> [2024-07-15].
- Eurydice (2024b). *Norway. 6.3 Assessment*.
<https://eurydice.eacea.ec.europa.eu/national-education-systems/norway/assessment-0> [2024-07-15].

- Eurydice (2024c). *Norway. 5.3 Assessment in single-structure education*. <https://eurydice.eacea.ec.europa.eu/national-education-systems/norway/assessment-single-structure-education> [2024-07-15].
- Eurydice (2024d). *Netherlands. 5.3 Assessment in primary education*. <https://eurydice.eacea.ec.europa.eu/national-education-systems/netherlands/assessment-primary-education> [2024-07-16].
- Gove, M. (2010). *The Democratic Intellect – what do we need to succeed in the 21st century?* Sir John Cass’s Foundation Lecture 2009. Sir John Cass’s Foundation.
- Government of Alberta (2010). *Guide to Education: ECS to Grade 12*. Alberta Education, Alberta.
- Government of Alberta (2023a). *Alberta Education Approved Literacy and Numeracy Assessments Grades 1–3*. Alberta Education, Alberta.
- Government of Alberta (2023b). *General Information Bulletin 2023–2024: Grades 6 and 9 Alberta Provincial Achievement Testing*. Alberta Education.
- Government of Alberta (2023c). *Guidelines for Interpreting Provincial Achievement Tests; Multiyear Reports*. <https://www.alberta.ca/system/files/educ-guidelines-for-interpreting-pat-multiyear-reports.pdf> [2024-07-17].
- Government of Alberta (2023d). *General Information Bulletin 2023-2024: Diploma Examinations Program*. https://www.alberta.ca/system/files/custom_downloaded_images/ed-diploma-exam-general-information-bulletin.pdf [2024-07-17].
- Government of Alberta (2023e). *Administering Diploma Exams*. <https://www.alberta.ca/administering-diploma-exams#toc-3> [2024-07-17].
- Government of Alberta (2024a). *Alberta’s School System*. <https://www.alberta.ca/albertas-school-system> [2024-07-17].
- Government of Alberta (2024b). *Student Learning Assessments*. <https://www.alberta.ca/student-learning-assessments> [2024-07-17].
- Government of Alberta (2024c). *Provincial Achievement Tests*. <https://www.alberta.ca/provincial-achievement-tests> [2024-07-17].

- Government of the Netherlands (u.å.). *Secondary school leaving examination*. <https://www.government.nl/topics/secondary-education/examination-in-secondary-education/secondary-school-leaving-examination> [2024-07-16].
- Government of the Netherlands (2023). *Request adapted final examination in secondary education*. <https://www.government.nl/topics/secondary-education/examination-in-secondary-education/request-adapted-final-examination-in-secondary-education> [2024-07-16].
- Graveleau, S. (2017). Parcoursup: ailleurs en Europe, quels critères pour accéder à l'université ? *Le Monde*, 2017-10-30. https://www.lemonde.fr/campus/article/2017/10/30/quels-sont-les-criteres-pour-acceder-a-l-universite-en-europe_5207848_4401467.html [2024-07-15].
- Haridus- ja Noorteamet [Education and Youth Board] (2021). *Final examinations for basic school*. Republic of Estonia. <https://harno.ee/en/final-examinations-basic-school> [2024-07-15].
- Haridus- ja Noorteamet [Education and Youth Board] (2024). *State examinations*. <https://harno.ee/en/state-examinations> [2024-07-15].
- Haridus- ja Teadusministeerium [Ministry of Education and Research] (2024). *Pre-school, basic and secondary education*. <https://www.hm.ee/en/education-research-and-youth-affairs/general-education/general-education-estonia> [2024-07-15].
- Higher Education Standards Panel (2016). *Improving the transparency of higher education education admissions*. Final report.
- Hong Kong Examinations and Assessment Authority (u.å.). *Territory-wide System Assessment (TSA)*. <https://www.bca.hkeaa.edu.hk/web/TSA/en/Introduction.html> [2024-08-06].
- Hong Kong Examinations and Assessment Authority (2018). *Grading Procedures and Standards-referenced Reporting in the HKDSE*. https://www.hkeaa.edu.hk/DocLibrary/Media/Leaflets/HKDSE_SRR_A4booklet_Mar2018.pdf [2024-08-06].

- Hong Kong Examinations and Assessment Authority (2021). *Marking Procedure for Written Papers*. https://www.hkeaa.edu.hk/en/HKDSE/assessment/markings/markings_procedure_for_written_papers/ [2024-08-06].
- Hong Kong Examinations and Assessment Authority (2023a). *Moderation of School-based Assessment Scores in the HKDSE*. https://www.hkeaa.edu.hk/DocLibrary/Media/Leaflets/moderation_of_SBA_score_en.pdf [2024-08-06].
- Hong Kong Examinations and Assessment Authority (2023b). *Hong Kong Diploma of Secondary Education Examination: Information on School-based Assessment*. https://www.hkeaa.edu.hk/DocLibrary/Media/Leaflets/SBA_booklet_eng.pdf [2024-08-07].
- International Baccalaureate Organization, IBO (2016). *Extended Essay. Guide*. https://xmltwo.ibo.org/publications/DP/Group0/d_0_eeyyy_gui_1602_1/files/Guide_Extended_essay_en.pdf [2024-08-07].
- International Baccalaureate Organization, IBO (2018). *Assessment principles and practices—Quality assessments in a digital age*.
- International Baccalaureate Organization, IBO (2022). *Access and inclusion policy*.
- International Baccalaureate Organization, IBO (2024). *IBO Sweden*. <https://www.ibo.org/country/SE/> [2024-08-07].
- infofinland.fi (2024). *Yrkeshögskolor*. <https://www.infofinland.fi/sv/education/universities-of-applied-sciences> [2024-07-15].
- IQB.a, Institut zur Qualitätsentwicklung im Bildungswesen (u.ä.). *Competence in Education*. https://www.iqb.hu-berlin.de/IQB_Flyer_EN_201.pdf [2024-07-17].
- IQB.b, Institut zur Qualitätsentwicklung im Bildungswesen (u.ä.). *Gemeinsame Abituraufgabenpools der Länder*. <https://www.iqb.hu-berlin.de/abitur> [2024-07-17].
- IQB, Institut zur Qualitätsentwicklung im Bildungswesen (2021a). *VERA – Ein Überblick*. <https://www.iqb.hu-berlin.de/vera/> [2024-11-01].

- IQB, Institut zur Qualitätsentwicklung im Bildungswesen (2021b). *VERA Online*. <https://www.iqb.hu-berlin.de/tba/online/> [2024-07-17].
- Johnes, R. (2016). The 11-plus and access to grammar schools. School performance, admissions, & capacity. *Education Policy Institute*. https://epi.org.uk/wp-content/uploads/2018/02/EPI_11_Plus_Examination_Analysis.pdf [2024-07-15].
- Kai, N.W. & Shafeeq, S. (2023). Teachers, students adapt to full subject-based banding that caters to diverse interests, strengths. *The Straits Times*, 2023-03-01. <https://www.straitstimes.com/singapore/politics/teachers-students-adapt-to-full-subject-based-banding-say-it-caters-to-diverse-interests-and-strengths> [2024-08-07].
- Kirby, P. (2016). *Shadow schooling. Private tuition and social mobility in the UK. Improving social mobility through education*. The Sutton Trust.
- Konguavi, T. (2019). *Teachers slam UCP plan to bring back Grade 3 standardized tests*. CBC, Radio-Canada. <https://www.cbc.ca/news/canada/edmonton/alberta-schools-teachers-ucp-standardized-tests-grade-3-1.5192668> [2024-07-17].
- Lichfield, J. (2015). French universities crisis: Low fees and selection lotteries create headaches in higher education. *The Independent*, 2015-09-25. <https://www.independent.co.uk/news/world/europe/french-universities-crisis-low-fees-and-selection-lotteries-create-headaches-in-higher-education-10517241.html> [2024-07-16].
- Meadows, M. (2016). How to choose your exam board. The Office of Qualifications and Examinations Regulation (Ofqual). *The Ofqual blog*. <https://ofqual.blog.gov.uk/2016/01/25/how-to-choose-your-exam-board/> [2024-07-15].
- Ministère de l'Éducation Nationale et de la Jeunesse (2013). *Organisation générale des études et les horaires des classes préparatoires scientifiques, accessibles aux titulaires d'un baccalauréat ou d'un titre admis en équivalence ou d'une dispense*. <https://www.education.gouv.fr/bo/13/Hebdo32/ESRS1317117A.htm> [2024-07-16].

- Ministère de l'Education Nationale et de la Jeunesse (2021).
Téléchargez le tableau de calcul de la note de bac en voie générale.
<https://www.education.gouv.fr/media/94490/download>
[2024-01-16].
- Ministère de l'Education Nationale et de la Jeunesse (2024a).
Examens et diplômés. <https://www.education.gouv.fr/examens-et-diplomes-41459> [2024-07-15].
- Ministère de l'Education Nationale et de la Jeunesse (2024b).
L'orientation en 3e et l'affectation en lycée.
<https://www.education.gouv.fr/reussir-au-lycee/l-orientation-en-3e-et-l-affectation-en-lycee-9257#edugouv-summary-item-7>
[2024-07-15].
- Ministère de l'Education Nationale et de la Jeunesse (2024c). *Le baccalauréat général.* <https://www.education.gouv.fr/reussir-au-lycee/le-baccalaureat-general-10457> [2024-07-15].
- Ministère de l'Education Nationale et de la Jeunesse (2024d).
Comment calculer votre note au baccalauréat.
<https://www.education.gouv.fr/reussir-au-lycee/comment-calculer-votre-note-au-baccalaureat-325511> [2024-07-15].
- Ministère de l'Enseignement supérieur et de la Recherche (2020).
Brevet de technicien supérieur (BTS).
<https://www.enseignementsup-recherche.gouv.fr/fr/brevet-de-technicien-superieur-bts-45915> [2024-07-16].
- Ministerium für Schule und Weiterbildung des Landes Nordrhein-Westfalen (2015). *School in North Rhine Westphalia: A quick and easy guide.* MSW 1/2015.
- Ministerium für Schule und Weiterbildung des Landes Nordrhein-Westfalen (2024). *Zentralabitur GOST: Fragen und Antworten.*
<https://www.standardsicherung.schulministerium.nrw.de/cms/zentralabitur-gost/fragen-und-antworten/> [2024-07-17].
- Ministry of Education Singapore (2018a). *Refreshing the holistic development profile (HDP). Annex B. I: 'Learn for Life' – Preparing Our Students to Excel Beyond Exam Results.*
<https://www.moe.gov.sg/-/media/files/news/press/2018/learn-for-life---annex-b.pdf> [2024-08-07].

- Ministry of Education Singapore (2018b). *'Learn for Life' – Preparing Our Students to Excel Beyond Exam Results*.
<https://www.moe.gov.sg/news/press-releases/20180928-learn-for-life-preparing-our-students-to-excel-beyond-exam-results> [2024-08-07].
- Ministry of Education Singapore (2023). *Autonomous universities*.
<https://www.moe.gov.sg/post-secondary/overview/autonomous-universities/> [2024-08-07].
- Ministry of Education Singapore (2024a). *Edusave account*.
<https://www.moe.gov.sg/financial-matters/edusave-account> [2024-08-07].
- Ministry of Education Singapore (2024b). *Edusave awards and scholarships*. <https://www.moe.gov.sg/financial-matters/awards-scholarships/edusave-awards> [2024-08-07].
- Ministry of Education Singapore. (2024c). *Curriculum for secondary schools*. <https://www.moe.gov.sg/secondary/schools-offering-full-sbb> [2024-08-07].
- Ministry of Education Singapore. (2024d). *Overview of Post-Secondary Education Institutions (PSEI)*.
<https://www.moe.gov.sg/post-secondary/overview> [2024-08-07].
- Nandhini, N. (2024). *Grading system in Canada: Ontario, Alberta, Quebec, Nova Scotia, Saskatchewan and British Columbia*.
<https://edvoy.com/articles/grading-system-in-canada/> [2024-10-29].
- National Center on Education and the Economy, NCEE (2024a). *Hong Kong*. Top performing countries.
<https://ncee.org/country/hong-kong/> [2024-08-06].
- National Center on Education and the Economy, NCEE (2024b). *Singapore*. Top performing countries.
<https://ncee.org/country/singapore/> [2024-08-07].
- New South Wales Education Standards Authority (u.å.). *Awarding grades*. <https://educationstandards.nsw.edu.au/wps/portal/nesa/k-10/understanding-the-curriculum/awarding-grades> [2024-08-07].
- New South Wales Government (u.å.). *Going to school*.
<https://education.nsw.gov.au/schooling/parents-and-carers/going-to-school> [2024-08-07].

- New South Wales Government (2023). *Reporting to parents*.
<https://education.nsw.gov.au/teaching-and-learning/curriculum/reporting-to-parents/report-requirements>
 [2024-08-07].
- New South Wales Government (2024a). *HSC*.
<https://www.nsw.gov.au/education-and-training/nesa/hsc>
 [2024-08-07].
- New South Wales Government (2024b). *Record of School Achievement (RoSA)*. <https://www.nsw.gov.au/education-and-training/nesa/leaving-school/rosa> [2024-08-07].
- OECD (2023a). How much time do students spend in the classroom? I: *Education at a Glance 2023: OECD Indicators*. Paris: OECD Publishing.
- OECD (2024a). *Public spending on education. Primary to post-secondary non-tertiary, % of GDP*. OECD.
<https://data.oecd.org/eduresource/public-spending-on-education.htm> [2024-12-12].
- OECD (u.å.a). 15-year-old students who agree (or strongly agree) with the statement "I feel like I belong at school". PISA 2018. I: *OECD Child Well-Being Data Portal*.
- OECD (u.å.b). Social inequality. I: *OECD, Better Life Index*.
- Ofqual (2017). *Inter-board comparability of grade standards in GCSEs, AS and A levels 2017*. Ofqual/17/6330.
- Parcoursup (2024). *Comment sont examinées les candidatures?*
<https://www.parcoursup.gouv.fr/candidater-sur-parcoursup/comment-sont-examinees-les-candidatures-1357>
 [2024-07-15].
- Petersson, O. (2010). *Studentexamen*. Pocketbiblioteket, nr 44. SNS förlag.
- Riigi Teataja (2010). *Basic Schools and Upper Secondary Schools Act*. Passed 09.06.2010. Translation published: 25.06.2014.
<https://www.riigiteataja.ee/en/eli/524092014014/consolide>
 [2024-07-15].
- Riigi Teataja (2011). *National curriculum for basic schools*. Passed 06.01.2011 No. 1. Translation published: 24.09.2014.

- <https://www.riigiteataja.ee/en/eli/524092014014/consolide>
[2024-07-15].
- Samordna opptak (2024a). *Universitet og høyskole – Kravkoder*.
<https://www.samordnaopptak.no/info/opptak/opptak-uhg/spesielle-opptakskrav/liste-over-kravkoder/index.html>
[2024-07-15].
- Samordna opptak (2024b). *Universitet og høyskole – Poengberegning*.
<https://www.samordnaopptak.no/info/opptak/opptak-uhg/poengberegning/index.html> [2024-07-15].
- Samordna opptak (2024c). *Fagskole – Poengberegning*.
<https://www.samordnaopptak.no/info/opptak/opptak-fsu/poengberegning/index.html> [2024-07-15].
- Singapore Examinations and Assessment Board (2021). *Primary School Leaving Examination (PSLE) 2021: Standard-referenced assessment*. <https://www.seab.gov.sg/docs/default-source/national-examinations/psle/pslemarkingprocess.pdf>
[2024-08-07].
- Singapore Examinations and Assessment Board (2024a). *PSLE*.
<https://www.seab.gov.sg/home/examinations/psle>. [2024-08-07].
- Singapore Examinations and Assessment Board (2024b). *2024 Primary School Leaving Examination. Examination rules and regulations for candidates*.
https://www.seab.gov.sg/docs/default-source/national-examinations/psle/psle_instructions_for_candidates.pdf
- Skolverket (2007). *Tio EU-länders utbildningssystem. En översiktlig bild*. Dnr 2007:2172.
- Skolverket (2016). *Betygssystem i internationell belysning*. Rapport framtagen av en fristående projektgrupp: Christian Lundahl, Magnus Hultén och Sverre Tveit.
- Skolverket (2020). *TIMSS 2019: Svenska grundskoleelevers kunskaper i matematik*.
- Skolverket (2023a). *Slutbetygen i grundskolan våren 2023*.
- Skolverket (2023b). *ICCS 2022: Kunskaper, attityder och engagemang i medborgar-, demokrati- och samhällsfrågor hos årskurs 8-elever i ett internationellt perspektiv*.

- Skolverket (2023c). *Grundskolan, Betyg årskurs 9, tabell 1C, läsåren 2018/19–2023/24.*
- SOU 2020:43. *Bygga, bedöma, betygssätta – betyg som bättre motsvarar elevernas kunskaper.*
- Standards and Testing Agency (2024). *National curriculum assessments: information for parents.*
<https://www.gov.uk/government/collections/national-curriculum-assessments-information-for-parents> [2024-07-15].
- Stromboni, C. (2018). Parcoursup: comment mettre en place une « sélection équitable » à l'université ? *Le Monde*, 2018-07-14.
https://www.lemonde.fr/campus/article/2018/06/14/parcoursup-comment-mettre-en-place-une-selection-equitable-a-l-universite_5314905_4401467.html [2024-07-15].
- Studentexamensnämnden (2023). *Vad är studentexamen?*
<https://www.ylioppilastutkinto.fi/sv/informationsservice/information-om-studentexamen/vad-studentexamen-ar> [2024-07-15].
- Studentexamensnämnden (2024a). *Poänggränser?*
<https://www.ylioppilastutkinto.fi/sv/utforandet-av-examen/poanggranser> [2024-07-15].
- Studentexamensnämnden (2024b). *Information till examinanderna.*
<https://www.ylioppilastutkinto.fi/sv/utforandet-av-examen/information-till-examinanderna> [2024-07-15].
- Studentexamensnämnden (2024c). *Examens struktur.*
<https://www.ylioppilastutkinto.fi/sv/utforandet-av-examen/examens-struktur> [2024-07-15].
- Studienet.dk (u.å.). *Guide till gymnasiet.*
<https://www.studienet.dk/guide-til-gymnasiet> [2024-11-06].
- Styrelsen for undervisning og kvalitet (2024). *Prøve- og eksamenskalender 2024/2025. Gymnasiale prøver og uddannelser. Børne- og Undervisningsministeriet.*
<https://www.uvm.dk/-/media/filer/uvm/udd/gym/pdf24/maj/240813-proeve-og-eksamenskalender-2024-2025.pdf> [2024-11-06].

- The Alberta Teachers' Association (2023). *Public education in Alberta continues world-class standing*. Media Release 5, December 2023. <https://teachers.ab.ca/news/public-education-alberta-continues-world-class-standing> [2024-07-17].
- Think Student (2022). *What Happens if You Miss an A-Level Exam?* <https://thinkstudent.co.uk/what-happens-if-you-miss-an-a-level-exam/> [2024-07-15].
- UCAS (2024). *UCAS Tariff Points*. <https://www.ucas.com/undergraduate/applying-university/entry-requirements/ucas-tariff-points> [2024-07-15].
- UddannelsesGuiden (2024a). *Eux. Med eux tager du både en erhvervsuddannelse og en gymnasial eksamen*. <https://www.ug.dk/uddannelser/artikleromuddannelser/omerhvervsuddannelser/eux> [2024-07-15].
- UddannelsesGuiden (2024b). *Fordeling af pladser på de gymnasiale uddannelser*. <https://www.ug.dk/6til10klasse/gymnasiale-uddannelserungdomsuddannelse/fordeling-af-pladser-paa-gymnasiale-uddannelser/fordeling-af-pladser-paa-de-gymnasiale-uddannelser> [2024-07-15].
- UddannelsesGuiden (2024c). *Alt om optagelse*. <https://www.ug.dk/videregaendeuddannelse/altomoptagelse> [2024-07-15].
- Universities Admissions Centre, UAC (2020). *Report on the scaling of the 2019 NSW higher school certificate*. <https://www.uac.edu.au/assets/documents/scaling-reports/scaling-report-2019-nsw-hsc.pdf> [2024-08-07].
- Universities and Colleges Admissions Service, UCAS (2019). *Singapore-Cambridge GCE A level*. <https://qips.ucas.com/qip/singapore-singapore-cambridge-gce-a-level> [2024-08-07].
- University of Alberta (2024). *Undergraduate Admissions & Programs. Competitive Requirements. Competitive Admission*. <https://www.ualberta.ca/admissions/how-to-apply/admission-requirements/competitive-requirements.html> [2024-07-17].

- UNESCO-UNEVOC (2022). *TVET Country Profiles: Poland*.
<https://unevoc.unesco.org/home/Dynamic+TVET+Country+Profiles/country=POL> [2024-07-16].
- UNICEF (2023). Table 15. Adolescents. I: *The State of the World's Children 2023: Statistical Tables*.
- Utbildningsstyrelsen (2019). *Grunderna för gymnasiets läroplan 2019. Föreskrifter och anvisningar 2019:2b*.
https://www.oph.fi/sites/default/files/documents/grunderna_for_gymnasiets_laroplan_2019.pdf
- Utbildningsstyrelsen (2020). *Bedömning av elevens lärande och kunskaper i den grundläggande utbildningen. Ändringar i grunderna för läroplanen för den grundläggande utbildningen 2014*. Utbildningsstyrelsen 10.2.2020.
- Utbildningsstyrelsen (2024a). *Bedömningskriterier för slutbedömningen i den grundläggande utbildningen*.
<https://www.oph.fi/sv/utbildning-och-examina/bedomningskriterier-slutbedomningen-i-den-grundlaggande-utbildningen> [2024-07-15].
- Utbildningsstyrelsen (2024b). *Gymnasieutbildningen*.
<https://www.oph.fi/sv/statistik/gymnasieutbildningen> [2024-07-16].
- Utdanningsdirektoratet (2020). *Kjennetegn på måloppnåelse – norsk 10. Trinn*. <https://www.udir.no/laring-og-trivsel/lareplanverket/kjennetegn/kjennetegn-pa-maloppnaelse-norsk-10.-trinn/> [2024-07-15].
- Utdanningsdirektoratet (2021). *Rammeverk for eksamen*.
<https://www.udir.no/eksamen-og-prover/eksamen/rammeverk-skriftlig-eksamen-i-lk20-og-lk20s/> [2024-07-15].
- Utdanningsdirektoratet (2022). *Vurdering av eksamen*.
<https://www.udir.no/eksamen-og-prover/eksamen/vurdering-av-eksamen/> [2024-07-15].
- Utdanningsdirektoratet (2023). *Kva er nasjonale prøver?*
<https://www.udir.no/eksamen-og-prover/prover/nasjonale-prover/om-nasjonale-prover/> [2024-07-15].
- Utdanningsdirektoratet (2024). *Administrere eksamen*.
<https://www.udir.no/eksamen-og-prover/eksamen/administrere-eksamen/> [2024-07-15].

- Wendt, H., Walzebug, A., Bos, W., Smith, D. S. & Bremerich-Vos, A. (2016) *Germany. PIRLS 2016 ENCYCLOPEDIA*. TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College.
- Yung, B. (2020). *Internal Assessments for Secondary School Places Allocation amid the Epidemic*. <https://www.edb.gov.hk/en/about-edb/press/insiderperspective/insiderperspective20200818.html> [2024-07-17].
- Øresunddirekt (2024). *Danska lärlingsutbildningar*. <https://www.oresunddirekt.se/se/jag-vill-studera-i-danmark/yrkesutbildning-och-larlingsutbildning/danska-larlingsutbildningar> [2024-07-15].

Statistiska analyser

I denna bilaga redovisas resultaten av de statistiska analyser som genomförts i utredningen och som ligger till grund för flera av de slutsatser och bedömningar som görs i betänkandets huvudtext.

1.1 Meritvärdeskalibrering – några exempel enligt en linjär modell

I detta avsnitt visas hur meritvärdeskalibrering går till utifrån några hypotetiska exempel. Dessutom visas hur betygsvärden för elever som inte deltagit på provet/proven kan kalibreras utan att påverka kalibreringen av övriga elevers betygsvärden.

1.1.1 Den linjära kalibreringsmodellen

$$BVK_{elev} = NSP_{skola} + (BV_{elev} - BV_{skola}) \cdot \left(\frac{\sigma_{NSP_{skola}}}{\sigma_{BV_{skola}}} \right) \quad (1)$$

BVK_{elev} = Elevens kalibrerade betygsvärde

NSP_{skola} = Skolans genomsnittliga resultat på slutproven

BV_{elev} = Elevens betygsvärde (ej kalibrerat)

BV_{skola} = Skolans genomsnittliga betygsvärde (ej kalibrerat)

$\sigma_{NSP_{skola}}$ = Spridningen i elevernas samlade provresultat på skolan

$\sigma_{BV_{skola}}$ = Spridningen i elevernas betygsvärde på skolan

Av ovanstående formel för den linjära kalibreringsmodellen framgår att kalibreringen av en elevs betygsvärde inte baseras på elevens individuella resultat på provet, om vi bortser från den indirekta och ringa effekt som elevens provresultat har på skolans medelvärde. Vad

som påverkar elevens kalibrerade betygsvärde är *skolans* genomsnittliga *provresultat*, det relativa avståndet mellan *elevens betygsvärde* och *skolans* genomsnittliga *betygsvärde* samt kvoten mellan *spridningen* i *skolans provresultat* och *spridningen* i *skolans betygsvärde*. I våra beräkningar används standardavvikelse (sd) som spridningsmått. Med standardavvikelsen menar vi ett mått på den genomsnittliga avvikelsen från medelvärdet av elevernas betygsvärden, provresultat respektive meritvärden. Ju större standardavvikelsen är, desto större är spridningen bland betygsvärden, provresultat respektive meritvärden.

$$MV_{elev} = 0,70 \cdot BVK_{elev} + 0,30 \cdot NSP_{elev}$$

MV_{elev} = Elevens meritvärde,

BVK_{elev} = Elevens kalibrerade betygsvärde,

NSP_{elev} = Elevens samlade nationella slutprovresultat.

Av ovanstående formel framgår hur en elevs meritvärde beräknas. Meritvärdet utgörs till 70 % av elevens kalibrerade betygsvärde och till 30 % av elevens individuella slutprovresultat.

1.1.2 Hypotetiska förenklade exempel

Äppelskolan

Det första exemplet, i tabell 1.1, gäller den hypotetiska Äppelskolan med totalt 11 elever, varav 10 har deltagit på provet. För att förenkla framställningen tänker vi oss ett enda prov och har valt heltalsvärden för elevernas betygsvärden trots att de i verkligheten sannolikt kommer att vara decimaltal eftersom de bygger på ett genomsnitt av ett flertal betyg på en 1–10 skala.

Tabell 1.1 Exempel 1: Kalibrering av meritvärden efter de genomsnittliga provresultaten för en hypotetisk skola, Äppelskolan

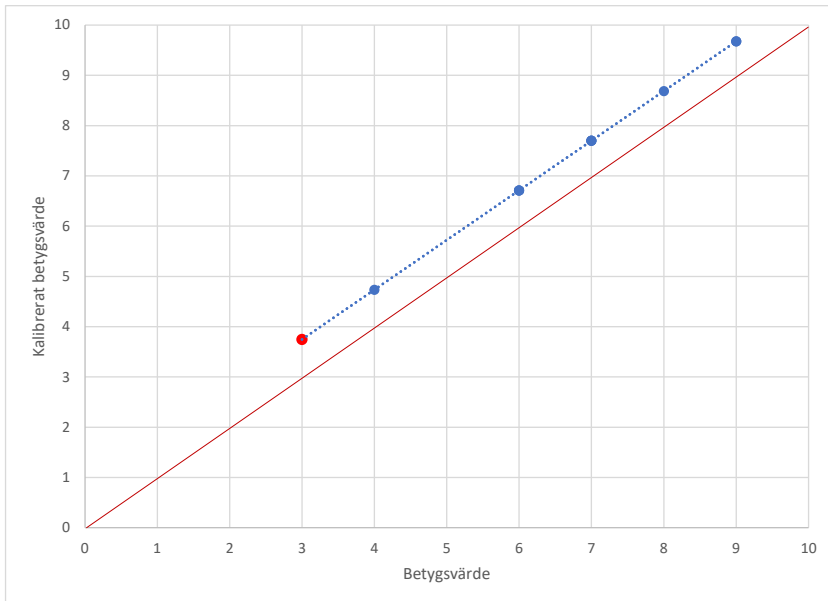
Elev	Betygsvärde	Provresultat	Kalibrerat betygsvärde	Meritvärde
1	9	10	9,7	9,8
2	9	9	9,7	9,5
3	8	9	8,7	8,8
4	7	8	7,7	7,8
5	7	8	7,7	7,8
6	7	7	7,7	7,5
7	6	7	6,7	6,8
8	6	7	6,7	6,8
9	6	6	6,7	6,5
10	4	5	4,7	4,8
11*	3	0	3,7	2,6
Medelvärde	6,90	7,60	7,60	7,60
Spridning (sd)	1,4	1,4	1,4	1,4

Anm.: Kalibreringsalgoritmen beräknas utifrån betygsvärden och provresultat för de elever som deltagit på proven. Elev nr 1–10 har deltagit på proven. Elev nr 11 har inte deltagit på proven men får ändå ett kalibrerat betygsvärde beräknat utifrån samma kalibreringsalgoritm som övriga elever på skolan.

* Bortfallslevers betygsvärde respektive kalibrerat betygsvärde ingår ej i det redovisade medelvärdet eller standardavvikelsen. Provresultatet för bortfallsleven sätts till 0 vid beräkning av dennes meritvärde.

Vi har förenklat exemplet genom att låta spridningen (mätt som standardavvikelsen) i betygsvärde respektive provresultat för samma 10 elever som deltagit på provet vara lika stor, 1,4 poäng för båda. Därmed reduceras den linjära modellen (1) till att bli en konstant positiv förskjutning med 0,7 poäng för samtliga elever på skolan. Denna transformation åskådliggörs i figur 1.1.

Figur 1.1 Grafisk representation av linjär transformation från betygsvärde till kalibrerat betygsvärde för exemplet Äppelskolan



Anm.: Vissa punkter i diagrammet representerar mer än en elev eftersom eleverna kan ha samma betygsvärde (se tabell 1.1). Den streckade linjen är referenslinje och representerar en 1:1 transformation. Den röda punkten representerar eleven som ej deltog på proven men som får ett kalibrerat betygsvärde enligt samma kalibreringsalgoritm som för övriga elever på skolan.

Elever som inte har deltagit på provet

För den elev (nr 11 i tabell 1.1) som inte deltagit på provet kan vi ändå kalibrera elevens betygsvärde, baserat på exakt samma kalibreringsalgoritm som övriga elever. Observera att bortfallselevens betygsvärde *inte* ingår i det medelvärde för skolan som ligger till grund för kalibreringen. Kalibreringsparametrarna ska beräknas enbart utifrån de elever som har deltagit på provet. Därefter *tillämpas* kalibreringen även på bortfallseleven genom intrapolering eller extrapolering. Detta innebär att även bortfallselevens kalibrerade betygsvärde (enligt detta exempel) blir 0,7 poäng högre än det okalibrerade betygsvärdet. Om bortfallseleven är hög- eller lågpresterande påverkar alltså inte hur övriga elevers betygsvärden kalibreras utan bestäms enbart utifrån relationen mellan betygsvärden och provresultat för de elever som deltagit på provet.

Päronskolan

Det andra exemplet gäller den hypotetiska Päronskolan. Av tabell 1.2 går att utläsa att Päronskolan har ett genomsnittligt betygsvärde på 7,8 och ett genomsnittligt provresultat på 6,3. Observera att dessa medelvärden är baserade endast på de 10 elever som deltagit på provet (elev nr 1–10). Eftersom provresultatet är lägre än betygsvärdet ska elevernas betygsvärde kalibreras ner med i genomsnitt 1,5 poäng.

Tabell 1.2 Exempel 2: Kalibrering av meritvärden efter de genomsnittliga provresultaten för en hypotetisk skola, Päronskolan

Elev	Betygsvärde	Provresultat	Kalibrerat betygsvärde	Meritvärde
1	10	10	9,96	9,97
2	9	8	8,29	8,21
3	9	8	8,29	8,21
4	9	8	8,29	8,21
5	8	7	6,63	6,74
6	8	6	6,63	6,44
7	7	6	4,97	5,28
8	6	5	3,31	3,82
9	6	3	3,31	3,22
10	6	2	3,31	2,92
11*	5	0*	1,65	1,15
Medelvärde	7,80	6,30	6,30	6,30
Spridning (sd)	1,4	2,3	2,3	2,3

Anm.: Kalibreringsalgoritmen beräknas endast utifrån betygsvärde och provresultat för de elever som deltagit på proven. Elev nr 1–10 har deltagit på proven. Elev nr 11 har inte deltagit på proven men får ändå ett kalibrerat betygsvärde beräknat efter samma kalibreringsalgoritm som övriga elever på skolan.

*Bortfallselevens betygsvärde respektive kalibrerade betygsvärde ingår ej i det redovisade medelvärdet och standardavvikelsen. Provresultatet för bortfallseleven sätts till 0 vid beräkning av dennes meritvärde.

I detta exempel skiljer sig spridningen i provresultat jämfört med spridningen i betygsvärde, mätt med standardavvikelsen. Standardavvikelsen är större i provresultat (2,3) jämfört med i betygsvärde (1,4). Därmed kommer kalibreringen inte bara att justera nivån utan även spridningen i betygsvärden. Detta åskådliggörs i figur 1.2 genom att den prickade linjen inte är parallell med den streckade referenslinjen (till skillnad från i föregående exempel). Lite förenklat kan spridningen mätas som avståndet mellan den högst respektive lägst presterande eleven av dem som deltagit på provet, projicerat dels på den

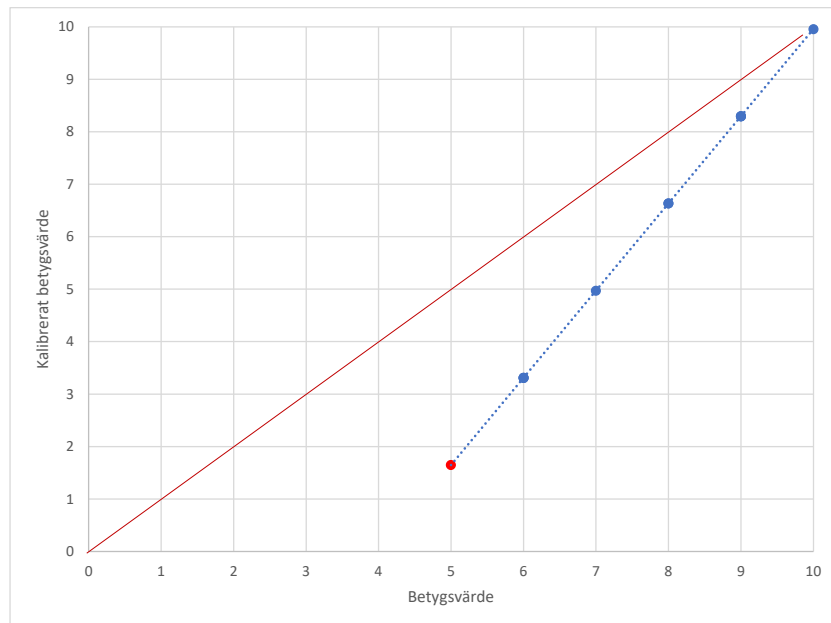
horisontella betygsvärdeaxeln, dels på den vertikala kalibrerade betygsvärdeaxeln. Avståndet i betygsvärde är 4 poäng medan avståndet i kalibrerat betygsvärde är 6,25 poäng.¹ Det som särskilt bör noteras här är att kalibreringen påverkar elever med relativt låga betygsvärden mer än elever med relativt höga betygsvärden. Det beror på att skolans betygssättning har varit mer generös för elever med låga betyg jämfört med elever med höga betyg, baserat på hur de presterat på proven. Denna ojämna justering är därmed helt i sin ordning då den justerar för att skolan systematiskt har varit mer generös i betygssättningen för de relativt lågpresterande eleverna.²

Även i detta exempel finns en elev som inte deltagit på provet men vars betygsvärde ändå kan kalibreras baserat på den kalibreringsalgoritm som beräknats utifrån de elever som deltagit på provet. Kalibreringen tillämpas även på bortfallseleven genom extrapolering i detta fall.

¹ Detta avstånd är bara en approximation på spridningen som mer exakt mäts med standardavvikelsen. Av den anledningen kommer inte avståndet i provresultat mellan den högst respektive lägst presterande eleven nödvändigtvis att vara lika med avståndet i kalibrerat betygsvärde. Däremot kommer standardavvikelsen i kalibrerat betygsvärde alltid att vara identiskt med standardavvikelsen i provresultat, vilket kan utläsas i figur 1.2.

² Kalibreringen tar inte hänsyn till den enskilda elevens provresultat. Att skolan systematiskt varit mer generös bland de lågpresterande eleverna framgår utifrån skolans genomsnittliga medelvärde och spridning i betygsvärde i relation till motsvarande medelvärde och spridning i provresultat.

Figur 1.2 Grafisk representation av linjär transformation från betygsvärde till kalibrerat betygsvärde för exemplet Pärönskolan



Anm.: Vissa punkter i diagrammet representerar mer än en elev eftersom eleverna kan ha samma betygsvärde (se tabell 1.2). Den streckade linjen är referenslinje och representerar en 1:1 transformation. Den röda punkten representerar elev nr 11 som inte deltagit på provet men vars betygsvärde kalibrerats baserat på samma kalibreringsalgoritm som för de elever som deltog på provet.

Kiwiskolan

Det tredje exemplet (tabell 1.3 och figur 1.3) baseras på simulerade men realistiska data där kalibreringsmodellen tillämpas i efterhand. Betygsvärdena baseras på grundskolans 16 ämnen och provresultatet på genomsnittet av fem nationella prov i matematik, engelska, svenska/svenska som andraspråk, NO och SO.³

Skolan består av 25 elever som alla har deltagit på samtliga fem nationella slutprov. Observera att i detta simulerade exempel har såväl betygsvärde, provresultat, kalibrerat betygsvärde samt meritvärde

³ Betygsvärdet och provresultaten i varje ämne har för hela årskullen årsvis standardiserats till ett nationellt medelvärde 6 och nationell standardavvikelse 1,5. Provresultaten i fysik, biologi och kemi har slagits ihop till ett genomsnittligt NO-provresultat och på samma sätt har geografi, historia, religion och samhällskunskap slagits ihop till ett genomsnittligt SO-provresultat. Därefter har ett samlat genomsnitt beräknats baserat på provresultaten i NO, SO, matematik, engelska och svenska.

transformerats till samma 1–10 poängskala som i de föregående exemplen trots att dagens betyg sätts på en annan skala.

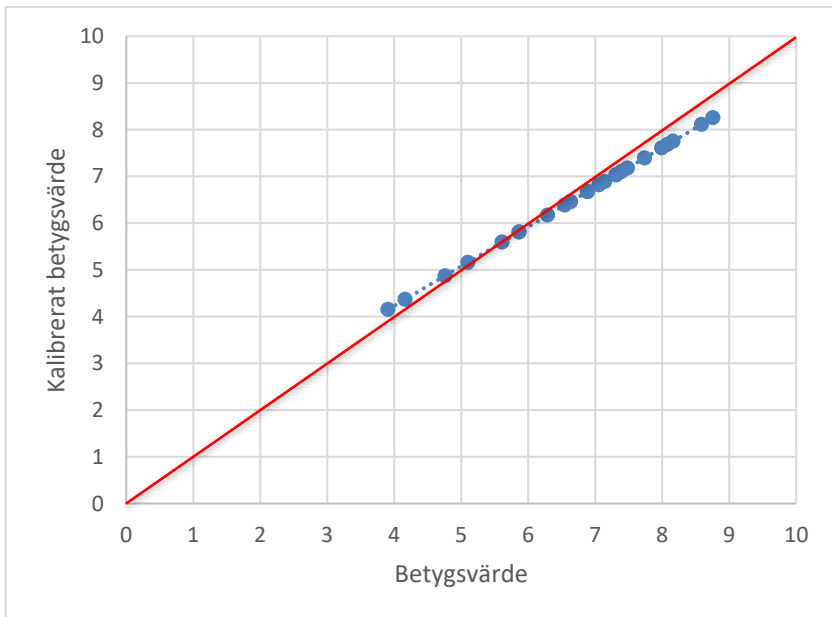
Här noteras att kalibreringen inte blir lika dramatisk som i exemplen med Pärönskolans och Äppelskolans. Skolans genomsnittliga betygsvärde är 6,73 och det genomsnittliga provresultatet är 6,54. Det innebär att betygsvärdet ska kalibreras ner med 0,19 poäng i genomsnitt. Dessutom ska spridningen minska något eftersom spridningen i samlat provresultat är något mindre än spridningen i betygsvärde (1,07 respektive 1,27).

Tabell 1.3 Exempel på hur den linjära kalibreringsmodellen appliceras på en simulerad skola, Kiwiskolan, åk 9

Elev	Betygsvärde	Provresultat	Kalibrerat betygsvärde	Meritvärde
1	8,76	7,77	8,26	8,11
2	8,59	8,80	8,11	8,32
3	8,16	8,06	7,75	7,85
4	8,08	7,74	7,68	7,70
5	7,99	7,28	7,61	7,51
6	7,74	7,80	7,39	7,51
7	7,48	7,52	7,18	7,28
8	7,40	7,01	7,11	7,08
9	7,31	5,98	7,03	6,72
10	7,31	6,78	7,03	6,96
11	7,14	7,29	6,89	7,01
12	7,06	6,48	6,82	6,72
13	6,89	5,98	6,67	6,47
14	6,89	6,24	6,67	6,54
15	6,63	6,01	6,46	6,32
16	6,63	6,75	6,46	6,54
17	6,55	6,55	6,39	6,43
18	6,29	5,99	6,17	6,12
19	5,87	5,95	5,81	5,85
20	5,87	6,01	5,81	5,87
21	5,61	5,46	5,59	5,55
22	5,10	5,19	5,16	5,17
23	4,76	5,74	4,87	5,13
24	4,16	3,91	4,37	4,23
25	3,91	5,18	4,15	4,46
Medelvärde	6,73	6,54	6,54	6,54
Spridning (sd)	1,27	1,07	1,07	1,05

Anm.: Utredningens egna beräkningar utifrån approximativ betygsstatistik och provresultat.

Figur 1.3 Grafisk representation av linjär transformation från betygsvärde till kalibrerat betygsvärde för Kiwiskolan



Anm.: Vissa punkter i diagrammet representerar mer än en elev eftersom eleverna kan ha samma betygsvärde (se tabell 1.3). Den streckade linjen är referenslinje och representerar en 1:1 transformation.

Det är intressant att observera att i detta exempel kommer inte alla elever att få ett lägre kalibrerat betygsvärde jämfört med det ursprungliga betygsvärdet, trots att skolans provgenomsnitt var lägre än genomsnittet av betygsvärdet. Det är en konsekvens av att betygsvärdena har för stor spridning i förhållande till vad provresultaten indikerar. Avståndet i poäng mellan de mest hög- och lågpresterande eleverna ska därför minskas något, vilket sker genom att de mest lågpresterande eleverna får ett något högre kalibrerat betygsvärde medan de mest högpresterande eleverna får den största sänkningen av betygsvärdet. Detta syns i figur 1.3 genom att de lägst presterande eleverna ligger ovanför den streckade referenslinjen medan flertalet elever ligger under referenslinjen.

Den interna rangordningen är densamma i betygsvärde och kalibrerat betygsvärde. Däremot är inte rangordningen nödvändigtvis densamma när vi jämför *meritvärdet* med det ursprungliga *betygsvärdet*, vilket kan utläsas i tabell 1.3. Där är eleverna rangordnade efter betygsvärdet, vilket alltså även motsvarar rangordningen i kalibrerat

betygsvärde. Elev 2 har således ett lägre *betygsvärde* och *kalibrerat betygsvärde* än elev 1 men ett högre *meritvärde* än elev 1, vilket är en konsekvens av att elevens *individuella* provresultat ges 30 % vikt i meritvärdet medan det kalibrerade betygsvärdet endast tar hänsyn till skolans *genomsnittliga* provresultat. Elev 2 har ett provresultat på 8,8 jämfört med 7,8 för elev 1. Om en elev presterar ett oväntat högt eller oväntat lågt provresultat så påverkar detta inte elevens kalibrerade betygsvärde mer än den marginella effekt det har på skolans genomsnittliga provresultat. Däremot kommer det att påverka elevens meritvärde eftersom det beräknas som det viktade genomsnittet av elevens kalibrerade betygsvärde (70 %) och elevens individuella provresultat (30 %).

På Kiwiskolan finns inga bortfallselever men om så hade varit fallet hade de inte påverkat kalibreringen eller det slutliga meritvärdet för övriga elever. Deltagande elevers slutliga meritvärde är oberoende av om eventuella bortfallselever är hög- eller lågpresterande. Det medför således inga fördelar att låta skolans mest lågpresterande elever stanna hemma och inte genomföra provet. Tvärtom finns inget att vinna på en sådan strategi för lågpresterande elever eftersom de därmed kommer att få noll (0) i provresultat när det slutliga meritvärdet ska beräknas.⁴

1.2 Konsekvenser av meritvärdeskalibrering i grundskolan

Kalibrering av meritvärdet i grundskolan kommer att innebära en delvis förändrad rangordning av elever jämfört med den rangordning som baseras på ett okalibrerat meritvärde och som representerar dagens betygssystem. I detta avsnitt redovisas resultaten från analyser av vilka konsekvenser ett kalibrerat meritvärde i grundskolan kan förväntas få på elevrankningen utifrån vissa centrala aspekter, däribland kön, migrationsbakgrund och socioekonomisk bakgrund. Dessutom analyseras konsekvenserna för den grupp elever som utgör de tio översta procenten i meritvärdefördelningen och som därmed i störst utsträckning kan antas konkurrera om att komma in på relativt attraktiva gymnasieprogram och skolor.

⁴ Värdet 0 gäller för de prov som eleven inte deltagit på men används endast vid beräkningen av elevens meritvärde. För kalibreringen tas inte elever med provresultat 0 med i beräkningen av skolans genomsnittliga resultat.

1.2.1 Data och metod

För att undersöka konsekvenserna av meritvärdeskalibrering har vi använt en kontrafaktisk ansats där vi med utgångspunkt i befintliga data, dvs. betyg, provbetyg och meritvärden, har applicerat den tänkta kalibreringsmodellen på meritvärdet.

För grundskolan, till skillnad från gymnasieskolan, bedöms analysen ge en mer robust indikation på konsekvenserna eftersom vi har tillgång till provresultat (provbetyg) från dagens fem nationella prov, dvs. lika många prov som meritvärdeskalibreringsmodellen är avsedd att använda. I gymnasieskolan har vi i nuläget endast tillgång till provresultat från fyra av de åtta prov som meritvärdeskalibreringen avser att använda.

Vi har dock inte tillgång till provpoängen från de nationella proven utan endast provbetygen, vilket innebär att en del information går förlorad. Dessutom är betyg och provbetyg satta efter en F–A-skala medan en kalibreringsmodell optimalt använder lärarsatta betyg på en 1–10 skala med 10 skalsteg samt provpoäng som standardiseras till motsvarande skala.

För grundskolan har vi, till skillnad från gymnasieskolan, tillgång till elevens samtliga ämnesbetyg, vilket innebär att vi inte nödvändigtvis behöver använda det färdigställda meritvärdet som baseras på en kodning där F = 0 poäng, E = 10 poäng, D = 12,5 poäng, C = 15 poäng, B = 17,5 poäng och A = 20 poäng. Vi har därför tagit fram ett alternativt meritvärde som vi benämner betygsvärde och som baseras på en 1–6 kodning av F–A-skalan där F = 1 och A = 6. Vi genomför alla analyser baserat på både befintligt meritvärde och vårt nya beräknade betygsvärde.

En annan svaghet med analysen av faktiska data är att lärarna redan har tagit hänsyn till de nationella provresultaten när de sätter betygen (i de ämnen som har nationella prov). Det innebär att de nationella provresultaten får en större betydelse än vad som är avsett i en meritvärdesmodell när den tillämpas på riktigt. Vi kan inte veta i vilken utsträckning lärarna har tagit hänsyn till elevernas provresultat när de satt betyg och det varierar med största sannolikhet beroende på lärare och skola samt även för olika elever trots att de har samma lärare. Detta bör man ha i åtanke när resultaten tolkas.

En sista svaghet med analysen är att lärarna rättar de nationella proven själva. Proven har alltså inte rättats centralt, vilket är en för-

utsättning för att meritvärdeskalibreringsmodellen ska kunna fungera på ett tillfredsställande sätt.

En viktig skillnad mot gymnasieanalysen är att vi för grundskolan delvis har kunnat inkludera elever som inte har deltagit på samtliga fem nationella prov men vars betyg ändå kan kalibreras.⁵

Sammantaget har vi kunnat applicera kalibreringen på årskohorterna 2013–2017, 2019 och 2022–2023.⁶ I genomsnitt har vi kunnat skatta kalibrerade meritvärden för cirka 93 % av alla elever som avslutade årskurs 9 under dessa år. I tabell 1.4 presenteras en sammanställning av de mest centrala uppgifterna.

Genom att årsvis standardisera samtliga provbetyg kan vi mäta alla resultat på samma skala. Först har vi tagit fram ett samlat provresultat för NO som ett genomsnitt av de tre årsvis standardiserade provbetygen i biologi, fysik och kemi. På samma sätt har ett samlat provresultat för SO tagits fram som ett genomsnitt av de årsvis standardiserade provbetygen i geografi, historia, religionskunskap och samhällskunskap. Slutligen har ett samlat totalt provresultat beräknats som genomsnittet av de årsvis standardiserade provresultaten i matematik, svenska/svenska som andraspråk, engelska, NO och SO.

⁵ Elever som inte har deltagit på minst ett prov kan omfattas av kalibreringen så länge de går i skolor där minst 50 % av eleverna har deltagit på samtliga fem nationella prov. Dessa elever kommer inte att påverka beräkningen av kalibreringen, men deras betyg kommer att kalibreras efter samma algoritm som övriga elever på skolan som deltagit på samtliga fem prov. Elever som inte deltagit på ett nationellt prov får i den slutliga sammanvägningen (av det kalibrerade betygsvärdet och elevens individuella provresultat) 0 poäng för de prov som eleven inte deltagit på. Om en elev t.ex. varit frånvarande på ett av fem prov kommer det genomsnittliga provresultatet från de fyra prov eleven har deltagit på att multipliceras med faktorn 4/5.

⁶ 2018 har exkluderats eftersom bortfallet på matematikprovet var över 60 % på grund av att provinnehållet spreds innan provtillfället. År 2020 och 2021 genomfördes inga nationella prov på grund av pandemin.

Tabell 1.4 Deskriptiv statistik för de data som analysen baseras på

Grundskolan årskurs 9, 2013–2017, 2019, 2022–2023

År		Löpnummer (max antal elever som det finns uppgifter för)	Meritvärde i dagens betygs- system	Betygsvärde (BV)	Nationellt slut- provsresultat (NSP)	Kalibrerat betygsvärde (BVK)	Meritvärde (MV)
2013	Antal	95 805	95 805	95 805	91 113	91 113	91 113
	Minimum		0,0	0,0	13,6	13,2	13,3
	Maximum		320,0	96,0	90,0	96,7	94,5
	Medelvärde		211,2	54,3	60,0	60,0	60,0
	Standardavvikelse		63,5	19,2	15,0	15,0	14,7
2014	Antal	96 559	96 559	96 559	92 754	92 754	92 754
	Minimum		0,0	0,0	14,8	13,6	13,9
	Maximum		320,0	96,0	88,9	96,6	94,0
	Medelvärde		212,8	55,0	60,0	60,0	60,0
	Standardavvikelse		65,0	19,6	15,0	15,0	14,7
2015	Antal	96 804	96 804	96 804	93 947	93 947	93 947
	Minimum		0,0	0,0	16,1	14,3	14,8
	Maximum		320,0	96,0	88,6	96,3	93,5
	Medelvärde		214,9	56,0	60,0	60,0	60,0
	Standardavvikelse		67,4	20,3	15,0	15,0	14,7
2016	Antal	103 853	103 853	103 853	99 725	99 725	99 725
	Minimum		0,0	0,0	21,6	18,1	19,2

År		Löpnnummer (max antal elever som det finns uppgifter för)	Meritvärde i dagens betygs- system	Betygsvärde (BV)	Nationellt slut- provsresultat (NSP)	Kalibrerat betygsvärde (BVK)	Meritvärde (MV)
	Maximum		320,0	96,0	84,8	94,2	90,8
	Medelvärde		211,6	55,7	60,0	60,0	60,0
	Standardavvikelse		76,5	22,1	15,0	15,0	14,7
2017	Antal	105 720	105 720	105 720	80 130	80 130	80 130
	Minimum		0,0	0,0	21,3	16,3	17,8
	Maximum		320,0	96,0	85,9	94,6	91,7
	Medelvärde		214,3	56,5	60,0	60,0	60,0
	Standardavvikelse		73,4	21,7	15,0	15,0	14,7
2019	Antal	112 202	112 202	112 202	103 671	103 671	103 671
	Minimum		0,0	0,0	18,7	15,3	16,3
	Maximum		320,0	96,0	86,1	93,4	90,7
	Medelvärde		218,6	57,9	60,0	60,0	60,0
	Standardavvikelse		71,9	21,5	15,0	15,0	14,7
2022	Antal	119 596	119 596	119 596	111 910	111 910	111 910
	Minimum		0,0	0,0	20,0	14,6	16,2
	Maximum		320,0	96,0	86,4	96,6	90,6
	Medelvärde		218,0	57,7	60,0	60,0	60,0
	Standardavvikelse		72,2	21,5	15,0	15,0	14,6
2023	Antal	119 855	119 855	119 855	115 502	115 502	115 502

År	Löpnummer (max antal elever som det finns uppgifter för)	Meritvärde i dagens betygs- system	Betygsvärde (BV)	Nationellt slut- provsresultat (NSP)	Kalibrerat betygsvärde (BVK)	Meritvärde (MV)
	Minimum	0,0	0,0	19,0	14,3	15,7
	Maximum	320,0	96,0	86,4	94,9	91,7
	Medelvärde	217,6	57,6	60,0	60,0	60,0
	Standardavvikelse	71,8	21,4	15,0	15,0	14,6
Totalt	Antal 850 394	850 394	850 394	788 752	788 752	788 752
	Minimum	0,0	0,0	13,6	13,2	13,3
	Maximum	320,0	96,0	90,0	96,7	94,5
	Medelvärde	215,1	56,4	60,0	60,0	60,0
	Standardavvikelse	70,6	21,0	15,0	15,0	14,6

Anm.: Betygsvärdet baseras på summan av de 16 betygen där F–A kodats som 1–6. Även provbetygen har kodats som 1–6 och som i denna analys utgör grunden för provresultaten. Nationellt slutprovsresultat (NSP) och kalibrerat betygsvärde (BVK) har årsvis standardiserats till medelvärde 60 och standardavvikelse 15.

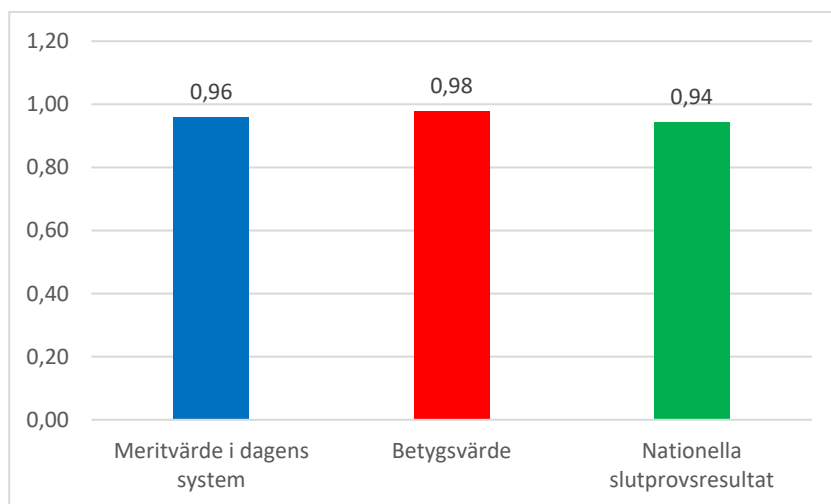
Meritvärde = kalibrerat betygsvärde (70 %) och elevens eget individuella samlade provresultat (30 %).

1.2.2 Konsekvenser för hela elevpopulationen och betygssystemet som helhet

Korrelationer och prognosvärden för det kalibrerade meritvärdet

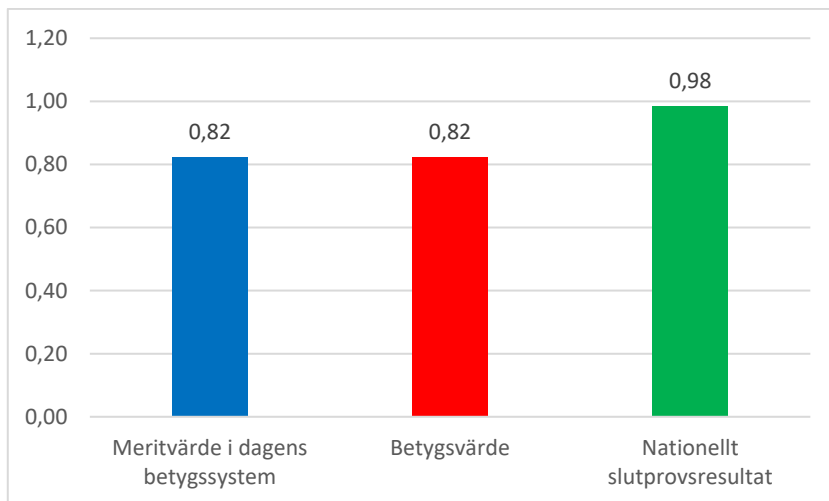
I figurerna 1.4 och 1.5 redovisas korrelationen mellan dagens meritvärde, betygsvärde resp. nationella slutprovresultat och det kalibrerade meritvärdet, vilket motsvarar det kalibrerade betygsvärdet plus elevens individuella provresultat som vägts in med 30 %. I figur 1.4 redovisas korrelationerna på elevnivå och i figur 1.5 på skolnivå.

Figur 1.4 Korrelation på elevnivå mellan det kalibrerade meritvärdet och andra utfallsmått från grundskolan



Källa: Utredningens egna beräkningar utifrån Skolverkets statistik.

Figur 1.5 Korrelation på skolnivå mellan det kalibrerade meritvärdet och andra utfallsmått från grundskolan



Källa: Utredningens egna beräkningar utifrån Skolverkets statistik.

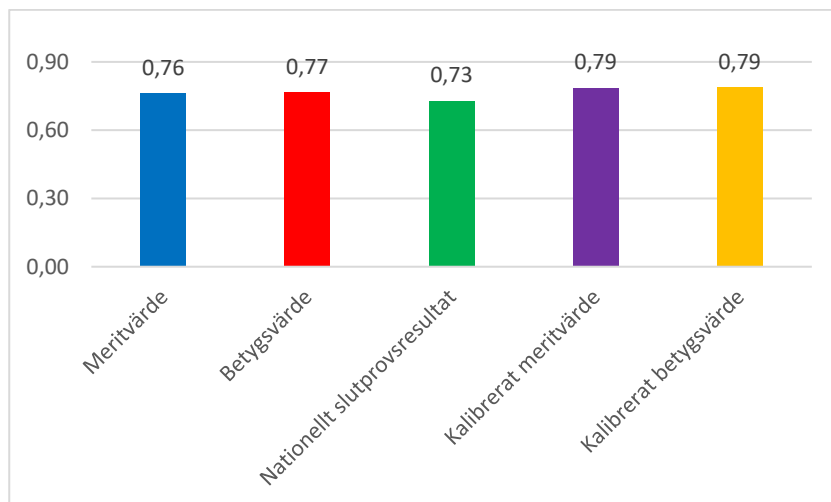
Om vi ser korrelationer nära 1, t.ex. 0,99, mellan det ursprungliga meritvärdet och det kalibrerat meritvärdet, skulle det innebära att förändringen i elevers rankning inte skulle förändras särskilt mycket och man kan ifrågasätta om den samhällsekonomiska nyttan av meritvärdeskalibrering överstiger kostnaderna. Om korrelationen i stället är mycket låg, t.ex. mindre än 0,6–0,7, kan det ifrågasättas om det kalibrerade betygsvärdet överhuvudtaget mäter samma typ av kunskaper som det ursprungliga meritvärdet.

På elevnivå ser vi relativt höga korrelationer mellan det kalibrerade betygsvärdet och motsvarande meritvärde respektive betygsvärde innan de kalibreras. Det kan tolkas som att konsekvenserna på systemnivå inte är alltför omfattande. Dessa korrelationer säger däremot ingenting om huruvida förändringarna är negativa eller positiva eller vilka som påverkas. Korrelationen med det samlade provresultatet är något lägre när det gäller elevnivå.

På skolnivå är korrelationerna lägre än på elevnivå för meritvärdet respektive betygsvärdet, vilket indikerar att skolornas genomsnittliga betygsvärden påverkas mer av kalibreringen än enskilda elever. Det inbördes mönstret är dock omvänt så att det nationella provresultatet på skolnivå uppvisar högre korrelation, 0,98, med det cali-

brerade betygsvärdet på skolnivå. Det är väntat eftersom kalibreringsmodellen utnyttjar de nationella provresultaten just på skolnivå.¹

Figur 1.6 Korrelation med kalibrerat meritvärde på gymnasieskolan och olika betygsmått från grundskolan, elevnivå



Anm.: Data gäller endast elever som avslutade årskurs 9 år 2015 och 2016 samt deltog på samtliga tre nationella prov i gymnasieskolan och dessutom gick på högskoleförberedande program. $N = 45\,240$ elever.
Källa: Utredningens egna beräkningar utifrån Skolverkets statistik.

I figur 1.6 visas korrelationer mellan olika betygsmått från grundskolan med det kalibrerade meritvärdet från gymnasieskolan, även kallat prognosvärde, dvs. hur väl t.ex. meritvärdet från grundskolan kan prediktera studieframgång i gymnasieskolan. Det ursprungliga meritvärdet uppvisar en korrelation på 0,76 med elevernas meritvärde från gymnasieskolan (egentligen jämförelsetal), vilket är högre än motsvarande korrelation mellan provresultaten och gymnasie-meritvärdet. Det kalibrerade betygsvärdet är dock högre, 0,79. Detta indikerar att kalibrering av meritvärdet/betygsvärdet leder till en

¹ Eftersom skolans genomsnittliga provresultat styr skolans kalibrerade betygsvärde och därmed även det kalibrerade meritvärdet borde korrelationen på skolnivå vara 1 och inte 0,98. Orsaken är att i de redovisade korrelationerna även ingår elever som *inte* deltog på proven och vilka då fått 0 som provresultat när deras kalibrerade meritvärde ska beräknas. Men dessa elever ingår ej när det genomsnittliga provresultatet och genomsnittliga betygsvärdet beräknas och som ligger till grund för kalibreringsalgoritmen för respektive skola. Därmed kommer korrelationen inte vara 1 utan något lägre. Hade vi beräknat korrelationen endast på de elever som deltog på proven hade korrelationen på skolnivå varit 1 mellan provresultat och kalibrerat meritvärde.

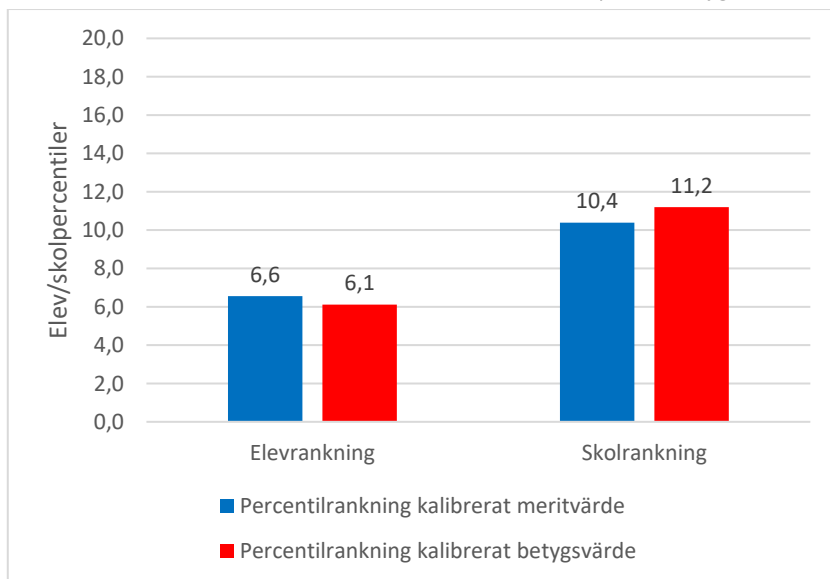
något bättre prognosförmåga samtidigt som det leder till en mer likvärdig betygssättning.

Förändringar i elevers och skolors ranking

I föregående avsnitt kunde vi med hjälp av korrelationen mellan ursprungligt meritvärde och kalibrerat betygsvärde få en uppfattning om hur stora förändringar som meritvärdeskalibrering kommer att få. Om korrelationen är 1,0 finns i praktiken ingen förändring. Men det är samtidigt svårt att tolka hur stor eller liten en förändring man kan vänta sig av en korrelation på 0,96. I figur 1.7 visas därför ett alternativt mått på förväntad förändring.

Figur 1.7 Genomsnittlig avvikelse i differensen i elev- respektive skolrankning

Beräknat som standardavvikelsen i skillnaden i rankningen mellan okalibrerade och kalibrerade meritvärden respektive betygsvärden



Anm.: Percentilrankning kalibrerat meritvärde = skillnad i percentilrankning mellan elevens/skolans ranking i kalibrerat meritvärde och det ursprungliga meritvärdet. Percentilrankning kalibrerat betygsvärde = skillnad i percentilrankning mellan elevens/skolans ranking i kalibrerat betygsvärde och det ursprungliga betygsvärdet (baserat på 1–6 kodning).

Källa: Utredningens egna beräkningar utifrån Skolverkets statistik.

I figur 1.7 framgår standardavvikelsen i differensen i elevens/skolans percentilrankning mellan det kalibrerade meritvärdet/betygsvärdet och det okalibrerade meritvärdet/betygsvärdet. Staplarna ska tolkas som den genomsnittliga förändring som elever respektive skolor kan förvänta sig i rankning, i absoluta tal.² Elever kan alltså i genomsnitt förvänta sig en förändrad percentilrankning på ungefär sex elevpercentiler (oavsett + eller –). För skolor kommer förändringen i skolrankning i genomsnitt att vara ungefär 11 skolpercentiler (+ eller –). Eftersom kalibreringen sker på skolnivå är det förväntat att effekterna blir större på skolnivå än på elevnivå. När skattade förändringar redovisas på detta sätt framstår de som mer betydande än när motsvarande korrelationer analyserades.

Förändring i skolskillnader

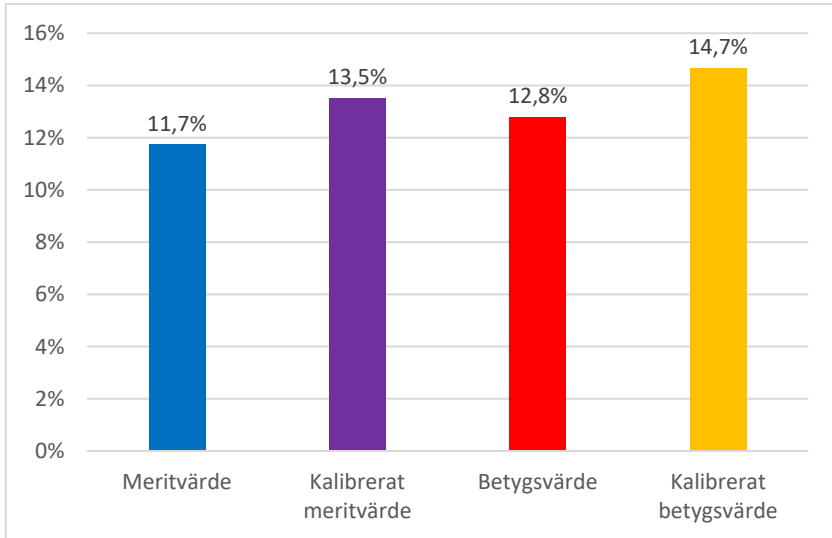
När meritvärdena kalibreras kan vi förvänta oss ökade skillnader mellan skolor eftersom 25 till 40 % av skolskillnaderna i betygssättning kan förklaras av s.k. relativ betygssättning, vilket innebär att betygssättningen tenderar att vara mer generös i lågpresterande skolor.³ I figur 1.8 visas den s.k. mellanskolvariationen som är ett mått på skolskillnader.⁴

² Genomsnittet i ändrad percentilrankning kommer alltid att vara noll när hela kohorten (som rankningen är baserad på) analyseras. Standardavvikelsen ger dock information om den genomsnittliga förändringen när vi bortser från om den är positiv eller negativ, alltså typ absolutvärde.

³ Skolverket (2019).

⁴ Mellanskolvariationen är andelen variation (variansen) i elevers resultat/betyg/meritvärde som kan förklaras med variation (variansen) mellan skolors genomsnittliga resultat/betyg/meritvärde.

Figur 1.8 Mellanskolvariation enligt meritvärde/betygsvärde respektive kalibrerat meritvärde/betygsvärde



Anm.: Meritvärdet i dagens betygssystem där F–A är kodat som 0–20. Kalibrerat meritvärde är baserat på samma kodning som meritvärde och elevens individuella provresultat har vägts in med 30 %. I betygsvärdet har F–A kodats som 1–6. Kalibrerat betygsvärde bygger på samma kodning som betygsvärde och där har elevens individuella provresultat vägts in med 30 %.

Källa: Utredningens egna beräkningar utifrån Skolverkets statistik.

Oavsett om vi mäter mellanskolvariationen med det ursprungliga okalibrerade meritvärdet eller det okalibrerade betygsvärdet visar figuren att mellanskolvariationen är större när vi mäter den med det kalibrerade meritvärdet/betygsvärdet. Ökningen uppgår till ungefär två procentenheter.

1.2.3 Konsekvenser för olika undergrupper av elever

I detta avsnitt redovisas den genomsnittliga förändringen i percentilrankning för undergrupper av elever som gått ut grundskolan. Förändringen beräknas som differensen i den percentilrankning som elevgruppen har enligt de kalibrerade betygsvärdena och de okalibrerade betygsvärdena.⁵ Förändringen är en faktisk förändring, dvs.

⁵ Vi redovisar inte skillnaden i kalibrerat meritvärde och det ursprungliga meritvärdet då mönstren är desamma. Det som skiljer betygsvärdet från meritvärdet är att betygsvärdet har beräknats utifrån att summera betyg som kodats med en 1–6 skala medan meritvärdet beräknats utifrån att summera betyg som kodats med en 0–20 skala (0; 10; 12,5; 15; 17,5; 20).

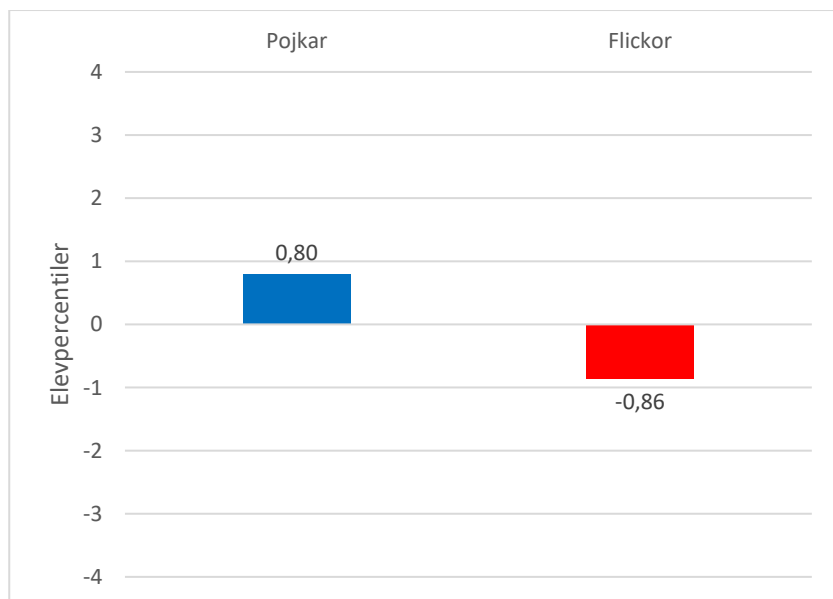
inte en prognos. Dessa analyser tar inte hänsyn till övriga bakgrundsfaktorer och ska därför inte tolkas som effekter. I efterföljande regressionsanalys tas hänsyn till övriga relevanta bakgrundsfaktorer och då kan resultaten tolkas som ett slags nettoeffekter.

Kön

Figur 1.9 visar att meritvärdeskalibrering kan förväntas leda till en genomsnittlig nedgång av flickors percentilrankning med en knapp elevpercentil jämfört med motsvarande percentilrankning enligt dagens betygssystem medan pojkarna kan förväntas få en uppgång i sin genomsnittliga percentilrankning med ungefär lika mycket.

Med dagens betygssystem ligger flickorna totalt sett i genomsnitt 12 elevpercentiler högre än pojkarna (56 mot 44). Med meritvärdeskalibrering skulle det gapet minska till drygt 10 elevpercentiler (55 mot 45).

Figur 1.9 Genomsnittlig förändring i percentilrankning av meritvärdeskalibrering för pojkar respektive flickor



Måttet är den genomsnittliga skillnaden i elevrankning mellan dagens okalibrerade meritvärde och det kalibrerade meritvärdet för respektive elevgrupp. Det kalibrerade meritvärdet utgörs av elevens kalibrerade betygsvärde (70 %) samt elevens individuella provresultat (30 %).

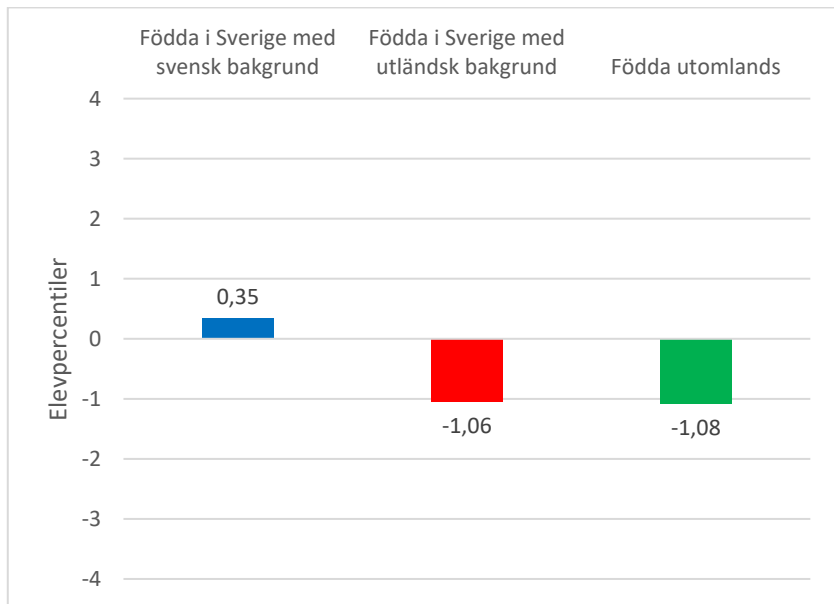
Källa: Utredningens egna beräkningar utifrån Skolverkets statistik.

Migrationsbakgrund

Figur 1.10 visar effekten av meritvärdeskalibrering för elevgrupper med olika migrationsbakgrund. Percentilrankningen för elever som är födda i Sverige med svensk bakgrund väntas i stort sett vara opåverkad med endast en marginell ökning. Elever som är födda i Sverige med utländsk bakgrund eller födda utomlands väntas få en försämrad percentilrankning med i genomsnitt cirka en elevpercentil.

Med dagens betygssystem är avståndet mellan elever födda i Sverige med svensk bakgrund och utlandsfödda elever drygt 19 elevpercentiler. Med meritvärdeskalibrering kan det avståndet förväntas öka till knappt 21 elevpercentiler.

Figur 1.10 Genomsnittlig förändring i percentilrankning av meritvärdeskalibrering för elever med olika migrationsbakgrund



Anm.: Måttet är den genomsnittliga skillnaden i elevrankning mellan dagens okalibrerade meritvärde och det kalibrerade meritvärdet för respektive elevgrupp. Det kalibrerade meritvärdet utgörs av elevens kalibrerade betygsvärde (70 %) samt elevens individuella provresultat (30 %).

Källa: Utredningens egna beräkningar utifrån Skolverkets statistik.

Föräldrars utbildningsnivå

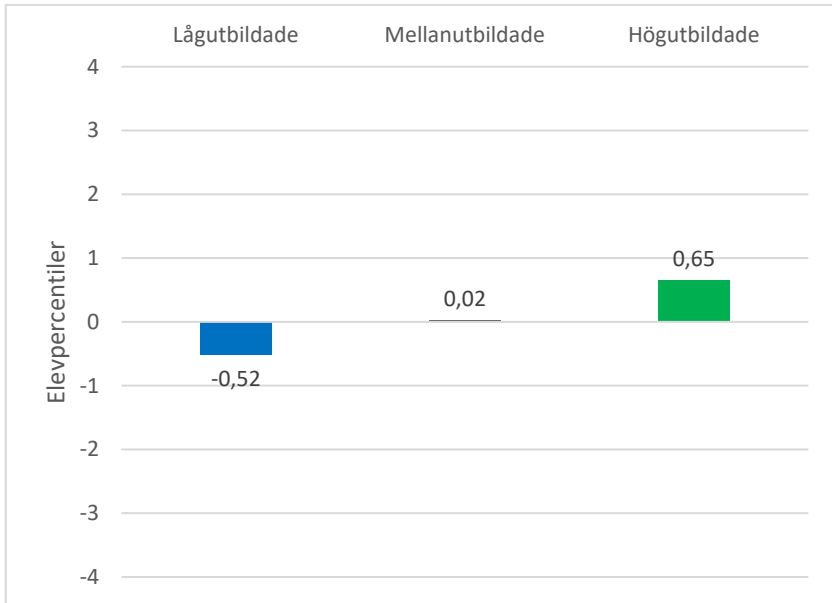
Måttet för föräldrars utbildningsnivå baseras på en kategorisering till tre kategorier som inte motsvarar den konventionella kategoriseringen med förgymnasial-, gymnasial- respektive eftergymnasial högsta utbildningsnivå och som ofta används i den officiella statistiken. I stället har vi här kategoriserat data för att få kategorier som inte är alltför små. Vi benämner kategorierna låg-, mellan- respektive hög utbildningsnivå. Andelen elever i respektive kategori i genomsnitt för alla år som ingår i analysen är ungefär 25, 50 respektive 25 %.⁶ Utfallsmåttet är liksom tidigare genomsnittet i förändrad rankning för respektive elevgrupp, vilket motsvarar skillnaden i percentilrankning mellan det kalibrerade betygsvärdet (där elevens individuella provresultat ingår med 30 %) och motsvarande percentilrankning i det okalibrerade betygsvärdet.

Figur 1.11 visar förändring i percentilrankning av meritvärdeskalibrering för elever med olika socioekonomisk bakgrund, mätt utifrån föräldrars utbildningsnivå. Elever vars föräldrar är kategoriserade som relativt lågutbildade kan förväntas försämra sin percentilrankning med en halv elevpercentil medan elever vars föräldrar kategoriseras som relativt högutbildade kan förväntas förbättra sin percentilrankning med ungefär lika mycket. Elever vars föräldrar kategoriseras som varken låg- eller högutbildade påverkas inte.

Skillnaden i genomsnittlig percentilrankning mellan elever med högutbildade och elever med lågutbildade föräldrar väntas öka från dagens 34 elevpercentiler till 35 (syns ej i figuren).

⁶ Utgångspunkten är en 7-gradig skala för respektive förälders högsta utbildningsnivå (SUN2000_old). Utifrån den har ett genomsnitt beräknats för varje elevs båda föräldrar, vilket innebär att skalan får 13 steg i stället för 7. Baserat på dessa 13 steg har sedan kategorierna definierats så att ungefär 25 % kategoriseras som *relativt* lågutbildade, 50 % som *relativt* mellanbildade samt 25 % som *relativt* högutbildade. Den lägsta kategorin (låg utbildningsnivå) motsvarar att båda föräldrarna som högst har en 2-årig gymnasieutbildning. Mellankategorin (medelhög utbildningsnivå) motsvarar att båda föräldrarna som högst har en 2-årig eftergymnasial utbildning. Den högsta kategorin (hög utbildningsnivå) motsvarar att båda föräldrarna minst har en 3-årig eftergymnasial utbildning. Eftersom kategoriseringen baseras på föräldrarnas genomsnittliga nivå på den 13-gradiga skalan kan olika kombinationer av föräldrarnas inbördes utbildningsnivå generera samma genomsnitt.

Figur 1.11 Genomsnittlig förändring i percentilrankning av meritvärdeskalibrering på elever vars föräldrar har olika utbildningsnivå



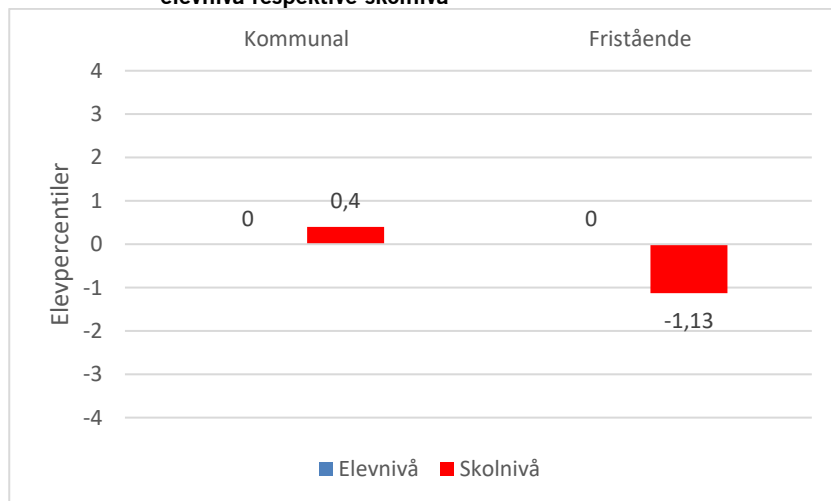
Anm.: Kategorierna baseras på en uppdelning så att cirka 25 % av eleverna hamnar i kategorin med relativt lågutbildade föräldrar, 50 % i mittenkategorin och 25 % i kategorin med relativt högt utbildade föräldrar.
 Källa: Utredningens egna beräkningar utifrån Skolverkets statistik.

Huvudman

Utfallsmåttet är liksom tidigare genomsnittet i förändrad rankning för respektive elevgrupp, vilket motsvarar skillnaden i percentilrankning mellan det kalibrerade betygsvärdet (där elevens individuella provresultat ingår med 30 %) och motsvarande percentilrankning i det okalibrerade betygsvärdet. Figur 1.12 visar den genomsnittliga förändringen i meritvärdeskalibreringen för elever som går på skolor med kommunal respektive fristående huvudman. Den skattade förändringen på elevnivå beräknas till noll. Det syns därmed inga svarta staplar i figuren.

På skolnivå, dvs. när vi analyserar förändring i skolors rankning, ser vi däremot att fristående skolor förväntas förlora i genomsnitt ungefär en percentilenhet i skolrankningen.⁷

Figur 1.12 Genomsnittlig förändring i percentilrankning av meritvärdeskalibrering på skolor beroende på huvudman, elevnivå respektive skolnivå



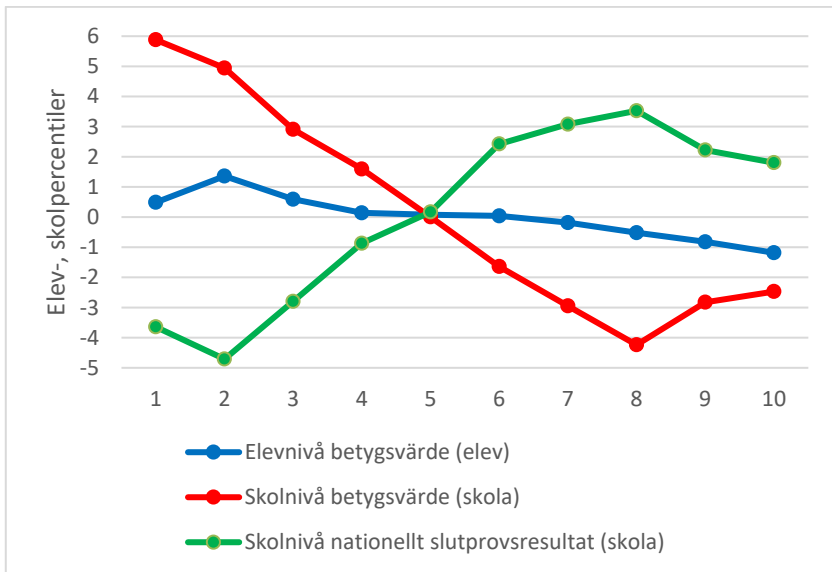
Källa: Utredningens egna beräkningar utifrån Skolverkets statistik.

Elever och skolor med olika betygsvärde och provresultat

Figur 1.13 visar flera olika resultat. Grunden för figuren är att elever och skolor har ordnats i tio lika stora grupper, s.k. deciler. Den blå linjen visar hur *elever* med olika relativ nivå på sina okalibrerade betygsvärden påverkas av kalibreringen. Elever med de 10 % lägsta betygsvärdena tillhör decil 1 och är längst till vänster på den horisontella axeln medan elever med de 10 % högsta betygsvärdena tillhör decil 10 och är placerade längst till höger på den horisontella axeln. Det framträder ett mönster på så sätt att elever med låga okalibrerade betygsvärden tenderar att förbättra sin percentilrankning med meritvärdeskalibrering medan det omvända blir följden för elever med höga okalibrerade betygsvärden. Effekterna är dock små, mindre än en elevpercentil för de flesta deciler.

⁷ Att de kommunala skolorna inte ser ut att vinna i skolrankning i samma utsträckning som friskolorna tappar i rankning beror på att de är många fler, vilket innebär att friskolornas nedgång fördelas på ett betydligt större antal kommunala skolor.

Figur 1.13 Genomsnittlig förändring i percentilrankning av meritvärdeskalibrering uppdelat på deciler av elevers respektive skolors betygsvärde samt skolors genomsnittliga provresultat



Anm.: Eleverna är uppdelade i 10 lika stora deciler där decil 1 innebär att elevernas tillhör de 10 % av elevkohorten med lägst betygsvärde och decil 10 innebär att eleverna tillhör de 10 % av elevkohorten med högst betygsvärde.

Källa: Utredningens egna beräkningar utifrån Skolverkets statistik.

Mönstret är detsamma men effekterna är betydligt större för motsvarande betygsvärden på skolnivå, dvs. skolornas genomsnittliga betygsvärde (röd linje). Skolor med de lägsta betygsvärdena (decil 1) väntas förbättra sin percentilrankning med sex skolpercentiler medan skolor som ligger på höga betygsvärden beräknas tappa i percentilrankning med 2–4 skolpercentiler.

Resultaten kan tolkas som att många av de skolor som uppvisar höga betygsvärden i genomsnitt tenderar att sätta glädjebetyg och när meritvärdena/betygsvärdena kalibreras tappar de i skolrankning.

En annan möjlig förklaring kan vara att högrepresterande skolor på något sätt missgynnas av själva meritvärdeskalibreringen. För att få klarhet i det har vi analyserat motsvarande förändring i skolrankning baserat på skolors resultat på de nationella proven.⁸ Skolorna är alltså

⁸ Observera att de provresultat vi har tillgång till i denna analys är rättade lokalt. Detta innebär att det kan finnas en viss skolbias även i provresultaten även om den är lägre än motsvarande

uppdelade på samma sätt i deciler som med betygsvärdena men nu efter de nationella provresultaten (grön linje). Nu ändras mönstret helt och hållet så att de skolor som tillhör de mest lågpresterande enligt proven (decil 1 och decil 2), förlorar i skolrankning med ungefär 4–5 skolpercentiler, samtidigt som de mest högpresterande skolorna (decil 8–10) förbättrar sin skolrankning med ungefär 2–4 skolpercentiler.

Detta resultat är fullt logiskt utifrån vetskapen om att *lokal relativ betygssättning* är vanligt förekommande i dagens betygssystem. Det innebär att, givet samma kunskap, är det svårare att få ett bra betyg på en skola med en hög andel högpresterande elever jämfört med på en skola med en hög andel lågpresterande elever.⁹ När meritvärdeskalibreringen tillämpas elimineras sådana effekter eftersom elevers kunskaper mäts på samma skala oavsett vilken skola de går på. Att mönstret är omvänt för skolor baserat på det okalibrerade betygsvärdet ger oss alltså en falsk bild av att själva kalibreringsmodellen skulle missgynna högpresterande skolor när det i själva verket är tvärtom. Skolor som sätter relativt generösa betyg kommer att förlora i skolrankning medan skolor som sätter relativt restriktiva betyg kommer att förbättra sin skolrankning jämfört med den rang de får med okalibrerade betyg.¹⁰

Högpresterande extremgrupper¹¹

Takeffekter kan uppstå om en skola har en relativt generös betygssättning, så att skolans genomsnitt kalibreras ner, men att de mest högpresterande eleverna slår i taket så att de, även i en skola med mindre generös betygssättning, skulle ha fått samma betygsvärde. Detta skulle kunna leda till att deras meritvärden kalibreras ner fastän de egentligen förtjänar det meritvärde de hade före kalibreringen. Figur 1.14 visar resultaten av analyserna om det finns en systematisk

skolbias i betygssättningen. De är troligt att de samvarierar så att om en skolas lärare i genomsnitt sätter generösa betyg så är det troligt att de också rättar proven generöst.

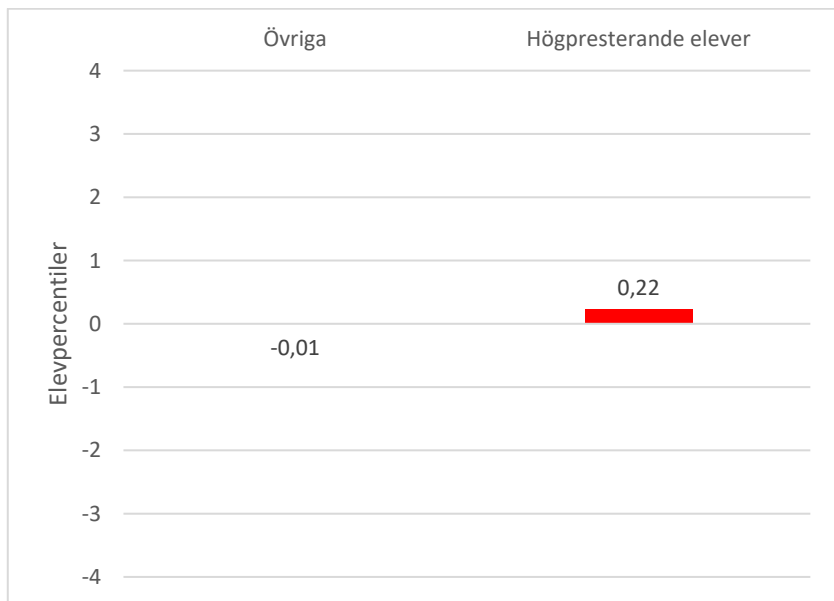
⁹ Skolverket (2019a).

¹⁰ Detta ska inte tolkas som att det i framtiden lönar sig att sätta mer restriktiva betyg och straffar sig att sätta relativt generösa betyg. Med meritvärdeskalibrering spelar det inte någon roll för elevernas meritvärden om skolan sätter relativt strikta eller relativt generösa betyg eftersom det kalibrerade betygsvärdet för skolan alltid kommer att bero på skolans genomsnittliga provresultat, oavsett skolans genomsnittliga betygsvärde innan kalibreringen.

¹¹ Vi har inte gjort motsvarande analys för lågpresterande extremgrupper, men i figur 1.13 syns att elever i de två lägsta decilerna, utifrån okalibrerat betygsvärde, inte förlorar i genomsnitt.

effekt på elever som är så pass högpresterande att de eventuellt kan påverkas negativt av takeffekter.

Figur 1.14 Genomsnittlig förändring i percentilrankning av meritvärdeskalibrering för högpresterande elever som inte underpresterar på de nationella proven



Anm.: Högpresterande elev = eleven har ett ostandardiserat betygsvärde på minst 90 poäng (motsvarar 10 A + 6 B där A–F kodats som 6–1) och presterar som sämst 3 elevpercentiler sämre på de nationella slutproven jämfört med motsvarande percentilrankning enligt betygsvärdet.

Källa: Utredningens egna beräkningar utifrån Skolverkets statistik.

I figur 1.14 särredovisas den elevgrupp som uppnår ett okalibrerat (ostandardiserat) betygsvärde på minst 90, vilket t.ex. motsvarar 10 ämnen med A i betyg och sex ämnen med B i betyg. För att räknas till denna grupp krävs dessutom att eleven, inte ”underpresterar” på de nationella proven, vilket definieras som att den inte får ha en percentilrankning på det samlade provresultatet som understiger motsvarande percentilrankning i betygsvärde med mer än högst tre procentenheter. Dessa elever utgör knappt tre procent av samtliga.

Enligt våra analyser påverkas inte denna elevgrupp negativt av meritvärdeskalibreringen, i stället förväntas de få en marginell förbättring med 0,2 elevpercentiler.

Analysens resultat är inte något slutgiltigt bevis för att det inte skulle kunna finnas systematiska negativa effekter på högpresterande elever som utsätts för takeffekter i samband med meritvärdeskalibrering. Detta är dock ett sätt att definiera och identifiera elevgruppen. Det bör också påpekas att denna typ av analys inte tar hänsyn till andra faktorer som kan påverka utfallet. I avsnitt 1.2.4 redovisar vi därför även resultat från en regressionsanalys där vi tar hänsyn till flera faktorer samtidigt. Den kan ge oss ytterligare svar på frågan.

Det ska slutligen klargöras att eventuella takeffekter inte beror på kalibreringsmodellen i sig utan är en konsekvens av att betygen på skolan inte satts på ett likvärdigt sätt. Detta eftersom det relativa avståndet mellan de högpresterande eleverna och övriga elever är för litet i förhållande till övriga relativa avstånd mellan olika elever inom samma skola. Om lärarna på skolan utnyttjar betygsintervallet på ett mer likvärdigt sätt så att skillnaderna i betygsvärden mellan olika elever korrekt reflekterar skillnaderna i kunskaper kommer inga takeffekter att uppstå.

Elever som inte deltar på samtliga nationella prov

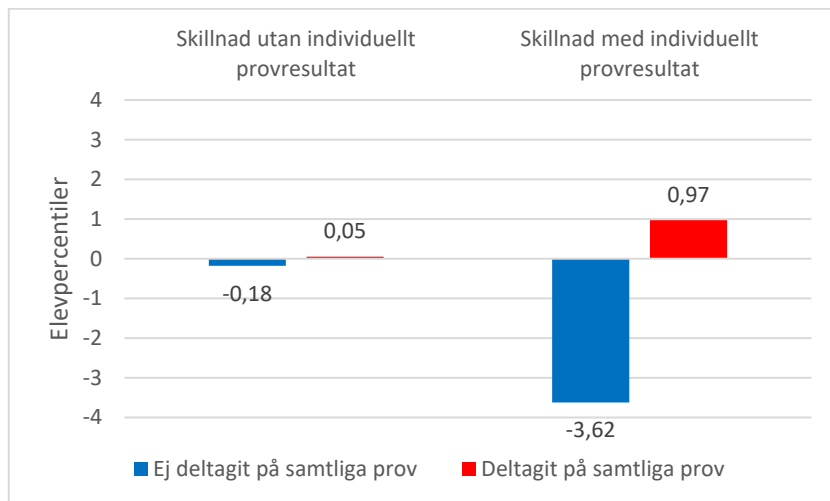
Så som meritvärdeskalibreringsmodellen fungerar beräknas kalibreringsparametrarna endast utifrån resultaten för de elever på skolan som har deltagit på samtliga fem nationella prov och endast utifrån dessa elevers betygsvärden. Om vi inte gjorde så skulle en skola där många högpresterande eller många lågpresterande elever inte deltog på proven leda till en orättvis kalibrering, åt ena eller andra hållet, för de elever som skrev provet.

Därför behöver vi hitta ett sätt att hantera de elever som inte deltar på alla prov. Det går att kalibrera dessa elevers betygsvärden men vi använder inte deras betygsvärdesuppgifter när vi beräknar kalibreringsparametrarna för skolan som helhet. Däremot tillämpar vi den beräknade kalibreringsalgoritmen även på de elever som inte deltagit på samtliga fem prov. I avsnitt 1.9 förklaras i detalj hur detta går till, vilka förutsättningarna är för att det ska fungera och varför det inte leder till orättvis kalibrering.¹² I detta avsnitt nöjer vi oss med

¹² Den begränsande faktorn för elevbortfall är framför allt om andelen frånvarande elever på en och samma skola är för hög.

att empiriskt se om det uppstår någon systematisk effekt som tyder på att kalibreringsmodellen inte fungerar som avsett.

Figur 1.15 Genomsnittlig förändring i percentilrankning av meritvärdeskalibrering på elever som deltagit respektive ej deltagit på samtliga nationella prov, elevnivå



Anm.: Skillnad utan individuellt provresultat = skillnad i elevers rankning mellan det kalibrerade betygsvärdet och det okalibrerade betygsvärdet utan att det individuella provresultatet har lagts till till det kalibrerade betygsvärdet. Skillnad med individuellt provresultat = skillnad i elevers rankning mellan det kalibrerade betygsvärdet, där elevens individuella provresultat vägts in med 30 %, och det okalibrerade betygsvärdet.

Källa: Utredningens egna beräkningar utifrån Skolverkets statistik.

I figur 1.15 visar det vänstra stapelparet den systematiska effekten av själva kalibreringen, *innan* elevens eget individuella provresultat har vägts in. Detta mått ger därför den mest relevanta insikten huruvida själva kalibreringen kan påverka elever som deltagit eller ej deltagit på olika sätt. Det tycktes inte finnas några systematiska effekter för varken den elevgrupp som deltagit på samtliga prov eller den elevgrupp som inte deltagit på samtliga prov.

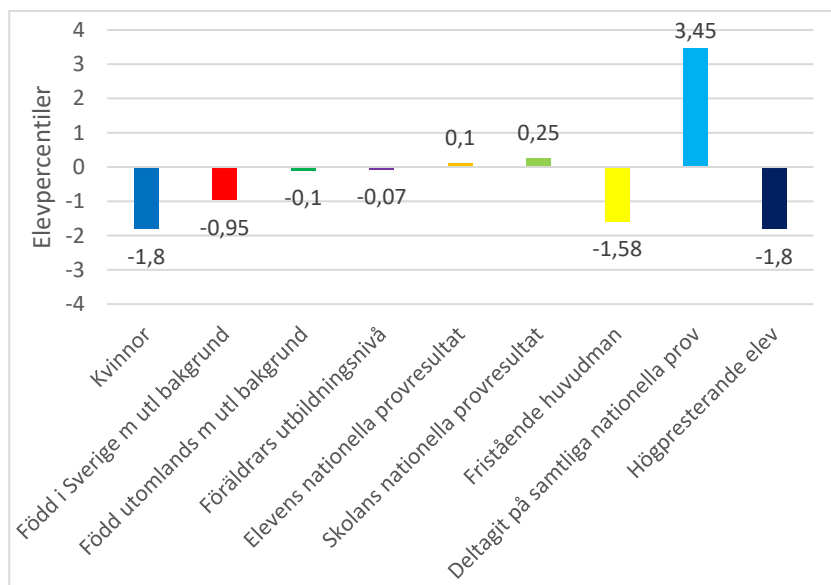
Det högra stapelparet visar effekten på det slutliga meritvärdet, dvs. där elevens individuella provresultat har 30 % vikt. Gruppen som *inte* deltagit på alla prov får resultatet 0 för varje prov som de inte deltagit på och detta påverkar naturligtvis dessa elevers slutliga meritvärde negativt. Vi ser också att effekten är $-3,6$ elevpercentiler, vilket är ett genomsnitt där en stor andel av eleverna bara missat ett prov medan några missat alla fem. Att vi ser en försämrad percentil-

rankning för denna elevgrupp är alltså helt logiskt och i linje med den beslutsregel som föreslås bli norm. Beslutsregeln ger incitament att delta på proven oavsett om man har höga eller låga lärarbetyg. Dessutom ges i det nya provsystemet möjlighet till ett uppsamlingsprov för dem som varit frånvarande vid det ordinarie provtillfället på grund av sjukdom. Dessa förändringar väntas medföra att betydligt fler genomför proven än i dag.

Slutsatsen är att modellen verkar fungera tillfredställande när vi tillämpar kalibrering även på de elever som inte deltagit på alla prov, under förutsättning att vi beräknar och tillämpar kalibreringsalgoritmen på ett korrekt sätt.¹³

1.2.4 Nettoeffekter på olika elevgrupper – regressionsanalys

Figur 1.16 Genomsnittlig förändring i percentilrankning av meritvärdeskalibrering baserat på regressionsanalys



Anm.: Standardiserade effektstorlekar. Samtliga effekter är statistiskt signifikanta på 5 % nivå. I regressionsanalysen skattas effekten på percentilrankning när hänsyn tas till alla övriga bakgrundsfaktorer som redovisas i figuren. Föräldrarnas utbildningsnivå mäts här med en 13-gradig approximativt kontinuerlig skala och inte med de tre kategorier som användes i figur 1.11.

Källa: Utredningens egna beräkningar utifrån Skolverkets statistik.

¹³ Eftersom denna applicering av meritvärdeskalibrering är baserad på historiska data där bortfallet (de som inte deltagit på proven) är högt, har vi valt att endast inkludera skolor där andelen elever som deltagit på alla prov är minst 50 %.

I figur 1.16 visas effekterna av meritvärdeskalibrering på olika elevgruppers percentilrankning när hänsyn tas till övriga elevgruppers påverkan; de kan tolkas som nettoeffekter.

Vi kan t.ex. se att effekten på flickors percentilrankning är $-1,8$ elevpercentiler, vilket är i linje med vad vi såg i figur 1.9, där skillnaden i effekt mellan pojkar och flickor uppgick till $1,7$ elevpercentiler, dvs. i praktiken ingen skillnad mot regressionseffekten. För utlandsfödda ser vi dock att nettoeffekten enligt regressionsanalysen endast uppgår till $-0,1$ elevpercentiler, dvs. praktiskt taget noll, men där effekten enligt figur 1.10 var $-1,1$ elevpercentiler. För elever som går på fristående skolor ökar i stället effekten från $0,0$ (figur 1.12) till $-1,6$ elevpercentiler. Det kan förklaras av att vi i regressionsanalysen har med skolans genomsnittliga provresultat som en förklarande variabel och den standardiserade effekten för skolans provresultat är $+0,25$ elevpercentiler.¹⁴

Denna variabel fångar upp effekten av att vi med hjälp av meritvärdeskalibrering eliminerar inslaget av relativ betygssättning. Det får konsekvenser för elever på fristående skolor eftersom dessa i genomsnitt, enligt (de lokalt rättade) provresultaten, presterar lite bättre än elever på kommunala skolor. Men när vi jämför fristående och kommunala skolor med samma prestationsnivå (samma genomsnittliga provresultat), vilket vi gör när vi inkluderar skolans genomsnittliga provresultat i regressionsmodellen, visar det sig att fristående skolor i genomsnitt har tillämpat en mer generös betygssättning än kommunala skolor. Effekten av detta är cirka $1,6$ elevpercentiler. När denna systematiska olikvärdighet elimineras får det en negativ effekt på $1,6$ elevpercentiler för elever på fristående skolor. När vi bara analyserade bruttoeffekten (figur 1.12) framkom inte denna effekt eftersom den relativa betygssättningen suddade ut detta mönster. När hänsyn tas till den relativa betygssättningen framträder däremot ett mer rättvisande mönster. Observera dock att denna effekt inte är särskilt stor och representerar ett genomsnitt av alla fristående skolor. Det finns kommunala skolor som sätter generösa betyg och det finns fristående skolor som sätter restriktiva betyg.

¹⁴ Att effekten är standardiserad innebär att när skolans genomsnittliga provresultat ökar med en standardavvikelse ökar elevens percentilrankning med $0,25$ standardavvikelse. De är därför inte helt jämförbara med de ostandardiserade effektstorlekarna men jämförbara med övriga standardiserade effektstorlekar t.ex. elevens provresultat eller föräldrars utbildningsnivå.

I våra analyser används en variabel som representerar de elever som har deltagit på samtliga fem nationella prov. De får en positiv effekt på 3,45 elevpercentiler relativt den elevgrupp som inte deltagit på samtliga nationella prov. Man kan även vända på det och säga att den genomsnittliga kostnaden för en elev av att inte delta på alla nationella prov är $-3,45$ elevpercentiler.

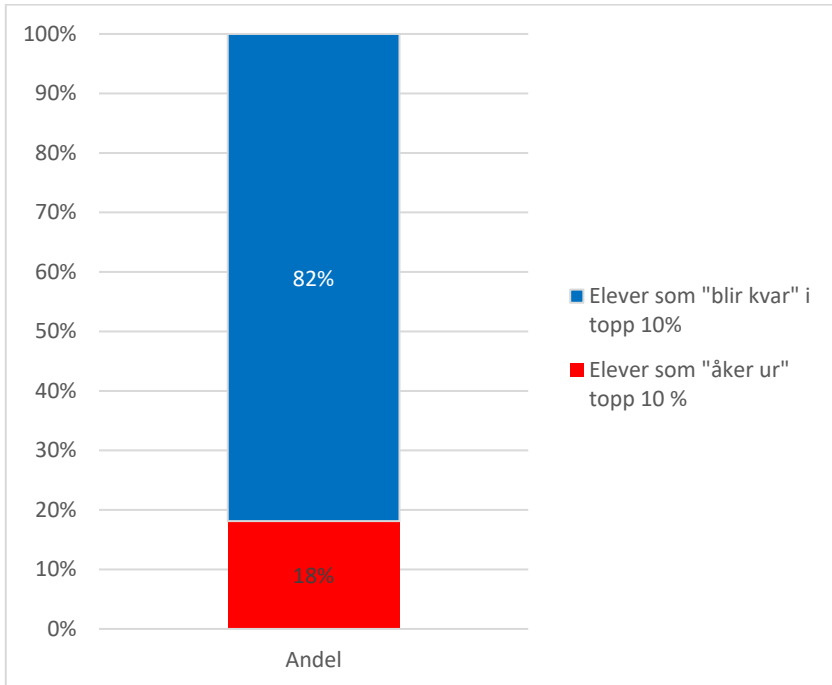
Att vi kontrollerar för huruvida eleverna deltagit på samtliga prov är viktigt när vi ska analysera nettoeffekten på den mest högpresterande elevgruppen, elever som har minst 90 poäng i betygsvärde av 96 möjliga (ej standardiserat och med betygskodning 1–6) och dessutom inte underpresterat mer än tre elevpercentiler på provet. Genom att vi nu i praktiken tar hänsyn till huruvida eleverna deltagit på samtliga prov så får vi en rättvisare jämförelse av vår extremgrupp. Det visar sig att de förlorar i percentilrankning med i genomsnitt 1,8 elevpercentiler, inte särskilt mycket men det kan ändå indikera att det finns vissa takeffekter som kalibreringsmodellen behöver ta hänsyn till innan den implementeras i betygssystemet.

Slutligen kan konstateras att nettoeffekten av föräldrars utbildningsnivå nu blir negativ från att ha varit positiv (figur 1.11). Den variabel som används i regressionsmodellen är dock approximativt kontinuerlig med 13 steg i stället för tre kategorier som i figur 1.11. I regressionsmodellen mäter vi därför om effekten av föräldrarnas utbildningsnivå är ökande oavsett var man befinner sig på skalan medan i figur 1.11 är det den genomsnittliga förändringen för elever med högutbildade föräldrar (utan hänsyn till övriga variabler).

Den sannolika huvudförklaringen till varför effekten byter tecken är att regressionsmodellen tar hänsyn till den relativa betygssättningen (genom att kontrollera för skolans genomsnittliga provresultat). Eftersom elever med högutbildade föräldrar dels tenderar att vara i genomsnitt något mer högpresterande, dels tenderar att gå på skolor som är högpresterande påverkas de negativt av den relativa betygssättningen. Med meritvärdeskalibreringen elimineras den relativa betygssättningen. I figur 1.11 har vi dock inte tagit hänsyn till detta explicit så den indirekta effekten fångas i stället till viss del upp av föräldrarnas utbildningsnivå. När vi i regressionsmodellen tar hänsyn till skolans genomsnittliga resultat på provet (och därmed effekten av att eliminera den relativa betygssättningen) finns inte längre någon positiv effekt av att ha högutbildade föräldrar på effekten på percentilrankning av meritvärdeskalibrering.

1.2.5 Förändringar i elevsammansättning för den högst presterande decilen (topp 10 %) av elevkohorten i grundskolan

Figur 1.17 Förändring i elevsammansättning före och efter meritvärdeskalibrering för den högsta decilen



Källa: Utredningens beräkningar utifrån Skolverkets statistik.

Figur 1.17 visar de elever som tillhör den högsta decilen, de 10 % mest högpresterande eleverna, enligt de lärarsatta betygen, dvs. det okalibrerade betygsvärdet. Dessa elever motsvarar således hela stapelns längd och utgör därför 100 %.

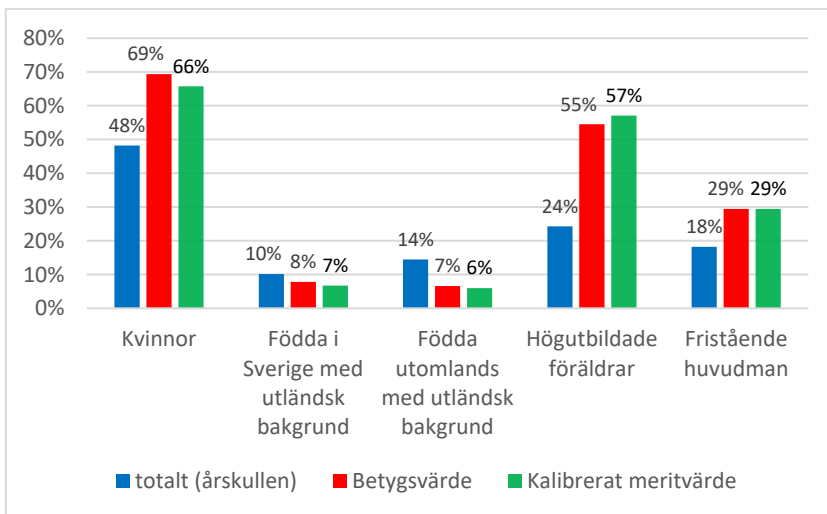
När vi tillämpar kalibrering på betygsvärdena och lägger till elevens eget individuella provresultat med 30 % vikt ser vi att elevrankningen förändras. Det medför att några av de elever som utifrån rankningen av de okalibrerade betygsvärdena tillhörde topp 10 % i hela elevkohorten, kommer att "åka ur" den gruppen när vi baserar rankningen på det kalibrerade betygsvärdet och väger in elevens individuella provresultat. Dessa elevers andel är rödmarkerad i stapeln och utgör 18 %

av hela den högpresterande elevgruppen. Resterande 82 %, representeras av den blå delen av stapeln som ”klarar sig kvar”, dvs. de tillhör den högpresterande elevgruppen oavsett om vi rangordnar efter det lärarsatta betygsvärdet eller det kalibrerade betygsvärdet.

Resultatet ger oss inblick i hur stora förändringar i toppen som vi kan förvänta oss i ett betygssystem där meritvärdeskalibrering tillämpas. I avsnitt 1.2 konstaterades att vi kunde vänta oss en *förändring* med 5–6 elevpercentiler för hela elevkohorten (se figur 1.13). Det måttet är inte helt jämförbart med det som visas i figur 1.17. Ingen elev kan bytas ut när vi analyserar hela elevkohorten (t.ex. en given årskohort) och den genomsnittliga förändringen i percentilrankning summerar alltid till noll. Däremot visar analysen i avsnitt 1.2 att det väntas ske förändringar i elevrankning som i genomsnitt innebär en förändring (både upp och ner) med 5–6 elevpercentiler. Man skulle kunna säga att av alla som åker ner i rankning så tappar de i genomsnitt 5–6 elevpercentiler och av alla som stiger i rankning så ökar de i genomsnitt 5–6 elevpercentiler.

När vi analyserar den högpresterande elevgruppen gör vi det på ett annat sätt eftersom det kan vara intressant att veta vilka som kan förväntas komma in, eller inte komma in, på en attraktiv gymnasieskola med hög konkurrens om utbildningsplatserna. Då ser vi att 18 % av alla elever kommer att bytas ut ”i toppen”. Det bedömer vi som en betydande förändring; införandet av meritvärdeskalibrering kan få viktiga effekter för den samhälleliga uppfattningen om rättvisa i fråga om tillträde till utbildning. Samtidigt ska vi komma ihåg att denna analys endast ska ses som en skattning av hur stora konsekvenserna kan bli. Vi har redan nämnt att de data som ligger till grund för dessa analyser inte kan anses helt jämförbara med de data som kommer att användas om meritvärdeskalibrering införs på riktigt. Till exempel har vi bara tillgång till provbetyg och inga provpoäng, vi använder de lärarsatta betygen som om de inte tagit hänsyn till elevens individuella provresultat och vi bortser dessutom från att proven är rättade av lärarna och inte centralt.

Figur 1.18 Andel av respektive elevgrupp i den högst presterande decilen (topp 10 %) av elevkohorten före respektive efter meritvärdeskalibreringen



Källa: Utredningens egna beräkningar utifrån Skolverkets statistik.

Figur 1.18 visar hur elevsammansättningen i den högpressterande elevgruppen väntas förändras utifrån några elevkategorier. Baserat på det okalibrerade betygsvärdet utgör flickor 69 % av den högpressterande elevgruppen (röd stapel i stapelgruppen längst till vänster i figuren) men den andelen sjunker till 66 % baserat på det kalibrerade betygsvärdet (grön stapel).

På liknande sätt minskar andelen elever som är födda i Sverige med utländsk bakgrund från 8 till 7 % och andelen utlandsfödda från 7 till 6 %. Samtidigt ökar andelen elever med högutbildade föräldrar från 55 till 57 % medan andelen elever från fristående skolor ligger kvar på 29 %. Motsvarande andelar för hela elevkohorten/årskullen (blå staplar) är 48 % flickor, 10 % födda i Sverige med utländsk bakgrund, 14 % födda utomlands, 24 % med högutbildade föräldrar och slutligen går 18 % i fristående skolor.¹⁵ Detta visar inte helt överraskande att elevsammansättningen bland högpressterande elever är snedvriden med en överrepresentation av flickor, elever med högut-

¹⁵ Att tillhöra elevgruppen med högutbildade föräldrar innebär här att föräldrarnas genomsnittliga utbildningsnivå på en 7-gradig skala (SUN2020) uppgår till minst 5,5 och är inte jämförbart med den mer konventionella definitionen där det räcker att minst en av föräldrarna ska ha en minst 3-årig eftergymnasial examen.

bildade föräldrar samt elever i fristående skolor och samtidigt en underrepresentation vad gäller elever med utländsk bakgrund. Meritvärdeskalibrering väntas fungera ”utjämnande” när det gäller andelen flickor och vara neutral vad gäller elever från fristående skolor. Där emot kommer den att öka både överrepresentationen vad gäller föräldrars utbildningsnivå och underrepresentationen av elever med utländsk bakgrund samtidigt som likvärdigheten ökar och rangordningen av elever blir mer rättvisande än i dag. Förändringarna bedöms sammantaget som relativt små.

Detta ger en bild av hur elevsammansättningen förändras i toppen av högpresterande elever. Det betyder inte nödvändigtvis att de elevgrupper, vars andel bland de högpresterande minskar, kommer att vara ”förlorare” när vi analyserar hela elevkohorten och med hänsyn till andra relevanta bakgrundsfaktorer.

1.3 Konsekvenser av meritvärdeskalibrering i gymnasieskolan

Kalibrering av meritvärden kommer att innebära en delvis förändrad rankning av elever jämfört med den rankning som baseras på det okalibrerade meritvärdet och som representerar dagens betygssystem. I detta avsnitt redovisas resultaten från analyser av vilka konsekvenser ett kalibrerat meritvärde kan förväntas få på elevrankningen i gymnasieskolan utifrån vissa centrala aspekter, däribland kön, migrationsbakgrund och socioekonomisk bakgrund. Dessutom analyseras konsekvenserna för den grupp elever som ligger i toppen av meritvärdesfördelningen och som konkurrerar om attraktiva utbildningar som bara en mindre andel av elevkohorten har en chans att komma in på som t.ex. läkarprogrammet och Handelshögskolan i Stockholm.

1.3.1 Data och metod

Att analysera konsekvenser av meritvärdeskalibrering för elever i gymnasieskolan är mer komplext än motsvarande analys för grundskolan som redovisades i avsnitt 1.2.

Vi har valt att begränsa data till de elever som tog gymnasieexamen 2018 respektive 2019 från de fem största högskoleförberedande

gymnasieprogrammen: Ekonomi (EK), Estetiskt (ES), Naturvetenskap (NA), Teknik (TE) och Samhällsvetenskap (SA). Dessutom begränsas urvalet av elever ytterligare genom att vi endast analyserar elever som har genomfört nationella prov i engelska 6, svenska eller svenska som andraspråk 3, matematik 1a/1b/1c samt högsta avslutad matematikkurs (t.ex. ma2b/ma2c/ma3b/ma3c/ma4). Slutligen inkluderas endast de elever som har skrivit högskoleprovet. Detta eftersom vi även vill analysera hur dessa resultat fördelar sig på de elever som kommer in respektive inte kommer in med de kalibrerade betygen (av de som kommer in med de okalibrerade betygen) för en viss intagningspoäng.

Dessa exkluderingskriterier leder till ett dataunderlag på 14 901 respektive 14 473 elever för 2018 respektive 2019. I tabell 1.5 ges en deskriptiv bild av data utifrån några bakgrundsvariabler.

Tabell 1.5 Deskriptiv bild av elevunderlaget som ingår i analysen på gymnasial nivå

Antal elever för vilka uppgifter finns tillgängliga för bakgrundsvariablerna i tabellhuvudet samt medelvärde och standardavvikelse för dessa

År		Jämförelsetal	Högskoleprov	Kön	Född utomlands	Föräldrars genomsnittliga utbildningsnivå
2018	Antal	14 901	14 901	14 901	14 901	14 861
	Medelvärde	15,592	0,788	0,550	0,100	4,605
	Standardavvikelse	2,357	0,350			1,190
2019	Antal	14 473	14 473	14 473	14 473	14 426
	Medelvärde	15,684	0,801	0,542	0,109	4,670
	Standardavvikelse	2,369	0,354			1,215
Totalt	Antal	29 374	29 374	29 374	29 374	29 287
	Medelvärde	15,638	0,794	0,546	0,105	4,637
	Standardavvikelse	2,363	0,352			1,203

Anm.: Variabeln Kön är kodad som 1 för kvinna och 0 för man. Medelvärdet anger då andelen kvinnor i dataunderlaget. Variabeln född utomlands är kodad som 1 om eleven är född utomlands, annars 0. Medelvärdet anger då andelen utlandsfödda i dataunderlaget. Föräldrarnas genomsnittliga utbildningsnivå mäts på en 7-gradig skala med 13 steg (dvs. 1, 1,5, 2... ..7). Standardavvikelse är ej redovisad för kön och född utomlands eftersom det saknar mening.

Källor: Skolverkets statistik och Umeå universitet.

För att kunna applicera meritvärdekalibrering i analyserna behövs nationella provresultat. Eftersom nationella prov endast genomförs i matematik, svenska och engelska är det dessa provresultat vi hänvisats till att använda. I en fullt implementerad modell kommer antalet prov att vara fler och även inkludera prov i natur- och samhällsorienterande ämnen.¹

Vi har också tvingats göra några förenklande antaganden. För det första har vi betraktat resultaten i matematik 1b och 1c, som om de vore samma prov, vilket de inte är även om de är relativt lika. Därmed kommer vi sannolikt att underskatta skillnaderna i matematikkunskaper mellan elever som läser på naturvetenskapliga programmet (NA) och teknikprogrammet (TE) jämfört med samhällsvetenskapliga programmet (SA), ekonomiprogrammet (EK) och estetiska programmet (ES).

Vi har i nästa steg utnyttjat de standardiserade resultaten från detta ”gemensamma” matematikprov 1 och identifierat den genomsnittliga resultatnivån för olika elevgrupper, beroende på vilken avancerad nivå i matematik som respektive elevgrupp läser (oavsett programtillhörighet). Dessa resultat används sedan för att placera provresultaten från respektive högre matematikkurs på samma skala som matematik 1-provet.

Låt oss t.ex. anta att eleverna som läser matematik 4, och därmed skriver nationellt prov i matematik 4, presterar 1,5 standardiserade poäng högre än genomsnittet för samtliga elever som genomförde matematik 1-provet. Då standardiseras resultaten från matematik 4-provet så att det genomsnittliga resultatet för matematik 4 ligger 1,5 poäng över det nationella genomsnittet för matematik 1-provet. På så sätt kan vi ta hänsyn till att matematik 4-provet är svårare än matematik 1-provet och placera båda proven på samma skala. Därmed straffas inte elever som läser en svår matematikkurs. Proven i matematik 2b, 3b och 3c skalas enligt samma princip. Därefter kan vi slå ihop respektive elevs två provresultat i matematik (grundläggande

¹ Samhällsvetare (EK, ES, SA) kommer, förutom grundläggande prov i matematik, avancerat prov i matematik, samt prov i engelska och svenska, även att genomföra prov i naturkunskap, samhällskunskap, religionskunskap och historia (totalt 8 prov som genererar 7 ämnesvisa provresultat som sedan slås ihop till ett samlat provresultat). Naturvetare (NA, TE) kommer, förutom grundläggande prov i matematik, avancerat prov i matematik, samt engelska och svenska, även att genomföra prov i SO, fysik, biologi och kemi (totalt 8 prov som genererar 7 ämnesvisa provresultat och som sedan slås ihop till ett samlat provresultat).

matematik och avancerad matematik) till ett enda jämförbart provresultat, trots att eleverna inte har skrivit samma avancerade prov.²

I svenska och svenska som andraspråk 3 har vi också förenklat och betraktat provresultaten i dessa ämnen som ett och samma provresultat. Samtliga elever läser och skriver nationellt prov i engelska 6. Där kan vi därför använda resultaten utan förenklingar eller begränsningar. Provresultaten i respektive ämne – matematik, svenska och engelska – standardiseras sedan årsvis och slås ihop till ett enda samlat provresultat vilket kan användas för kalibrering.

Sammantaget innebär ovanstående att de resultat som redovisas i detta avsnitt måste beaktas med viss försiktighet eftersom de inte fullt ut reflekterar hur en implementerad meritvärdesmodell skulle falla ut.

1.3.2 Kalibreringssteget

Vi har i denna analys för gymnasieskolan, på samma sätt som för grundskolan, tillämpat en linjär kalibreringsmodell:

$$BVK_{elev} = NSP_{skola} + (BV_{elev} - BV_{skola}) * \left(\frac{\sigma_{NSP_{skola}}}{\sigma_{BV_{skola}}} \right)$$

BVK_{elev} = Elevens kalibrerade betygsvärde

NSP_{skola} = Skolans genomsnittliga resultat på slutproven

BV_{elev} = Elevens betygsvärde (ej kalibrerat)

BV_{skola} = Skolans genomsnittliga betygsvärde

$\sigma_{NSP_{skola}}$ = Standardavvikelsen i elevernas betygsvärde på skolan

$\sigma_{BV_{skola}}$ = Standardavvikelsen i elevernas samlade provresultat på skolan.

Skolan som enhet ska för gymnasieskolan tolkas som program-skola-enhet. Det innebär att alla elever som går på samma program på samma skola betraktas som den grupp på vilken kalibreringen baseras.

² Genom denna metod lägger vi matematikproven i matematik 2b, 3b, 3c och 4 på samma skala som matematik 1-provet, genom att utnyttja det för alla elever gemensamma matematik 1-provet som ankare. Detta innebär att resultaten för de olika proven blir jämförbara.

1.3.3 Meritvärdesberäkning

I det sista steget beräknas elevens meritvärde som ett vägt genomsnitt av elevens betygsvärde (70 %) och elevens individuella provresultat (30 %).³ Det ger oss ett ”nytt” kalibrerat meritvärde, vilket kan jämföras med det ursprungliga meritvärdet (jämförelsetalet).

Högskoleprovskalibrerat meritvärde

Enligt samma princip som vi kalibrerar meritvärdena med hjälp av de nationella proven har vi i analyserna även kalibrerat meritvärdena baserat på elevernas resultat på högskoleprovet.⁴ I den slutliga meritvärdesberäkningen i dessa analyser har vi också använt elevens individuella resultat från högskoleprovet på samma sätt som de nationella provresultaten.

1.3.4 Konsekvenser för hela elevpopulationen och betygssystemet som helhet

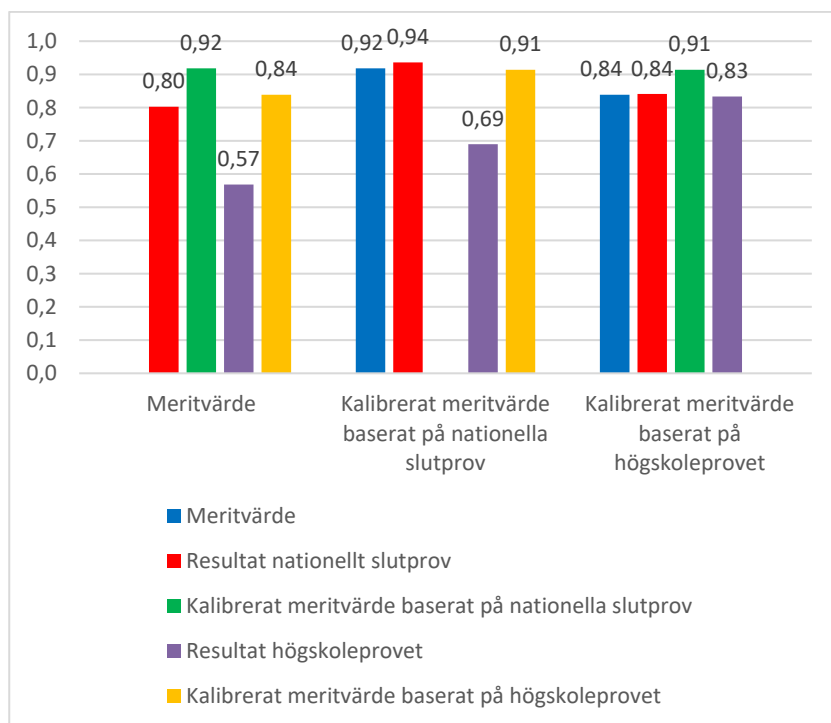
Korrelation mellan kalibrerat och okalibrerat meritvärde

Ett första mått som är relevant att undersöka är hur väl de två alternativa kalibrerade meritvärdena samvarierar med originalmeritvärdena. Korrelationen kan variera mellan $-1,0$ och $+1,0$ där ett positivt tal anger en positiv samvariation och en negativ korrelation en negativ samvariation. Om korrelationen är mycket hög, säg $0,99$, blir det ganska meningslöst att genomföra reformen. Om korrelationen i stället är för låg, säg mindre än $0,8$, kan det ifrågasättas om det kalibrerade meritvärdet överhuvudtaget mäter samma grundläggande kunskaper som betyget avser att mäta.

³ Se kapitel 5 i betänkandets volym 1 för motivering till val av vikter.

⁴ Data från Umeå universitet.

Figur 1.19 Korrelation mellan de percentilrankade kalibrerade respektive okalibrerade meritvärdena samt nationella prov- respektive högskoleprovsresultat, elevnivå



Anm.: Samtliga variabler är percentilrankade. Meritvärde = okalibrerat percentilrankat meritvärde (jämförelsetal). Resultat nationellt slutprov: percentilrankat. Kalibrerat meritvärde baserat på nationella prov: percentilrankat och här ingår elevens individuella provresultat med 30 %. Resultat på högskoleprovet: percentilrankat. Kalibrerat meritvärde baserat på högskoleprovet: percentilrankat och här ingår även elevens individuella högskoleprovsresultat med 30 %.

Källor: Utredningens egna beräkningar utifrån Skolverkets och Umeå universitets statistik.

I figur 1.19 framgår att korrelationen mellan det percentilrankade meritvärdet och ett meritvärde kalibrerat med nationella provresultat (NP-kalibrerat meritvärde) ligger på 0,92, medan motsvarande korrelation med ett meritvärde kalibrerat med högskoleprovsresultat (HP-kalibrerat meritvärde) är 0,84. Korrelationen mellan de två kalibrerade meritvärdena ligger på 0,91. Resultaten framstår som rimliga.

Korrelationen mellan det ursprungliga meritvärdet och det nationella slutprovet är 0,80 medan motsvarande korrelation med högskoleprovet är betydligt lägre, 0,57. Denna förhållandevis låga korrelation indikerar att högskoleprovet inte mäter samma kunskaper

som betygen avser att mäta och/eller faktiskt mäter. En annan möjlig förklaring eller delförklaring till den låga korrelationen kan vara att betygen inte är likvärdiga. Då blir det intressant att konstatera att korrelationen med det NP-kalibrerade meritvärdet, som alltså till viss del ska hantera den olikvärdiga betygssättningen, är högre, 0,69.

Utifrån det faktum att högskoleprovet är en oberoende mätning av elevens kunskaper kan denna ökning med 12 procentenheter tolkas som att det NP-kalibrerade meritvärdet är mer likvärdigt än det ursprungliga. En mindre välvillig tolkning är att det NP-kalibrerade meritvärdet bara mäter mer av samma kunskaper som högskoleprovet mäter, utan att likvärdigheten har förbättrats. Den första tolkningen framstår dock som mer rimlig även om den inte kan bevisas eftersom vi saknar uppgift om elevens ”sanna kunskap”.⁵ Observera att vi inte kan använda samma argumentation när det gäller det HP-kalibrerade meritvärdet. För det HP-kalibrerade meritvärdet är korrelationen med det percentilrankade resultatet på högskoleprovet visserligen högre, 0,83, men högskoleprovet är inte en oberoende mätning av elevens kunskaper i detta fall eftersom det också använts för att kalibrera meritvärdet i denna modell.

För att få ett jämförande perspektiv på dessa svårtolkade korrelationer kan nämnas att korrelationen mellan provbetyg och kursbetyg i svenska 3, engelska 6 och matematik 1 (vilka tillsammans med matematik 2 ligger till grund för det samlade provresultat som används för kalibrering) ligger på 0,82, 0,89 respektive 0,94.

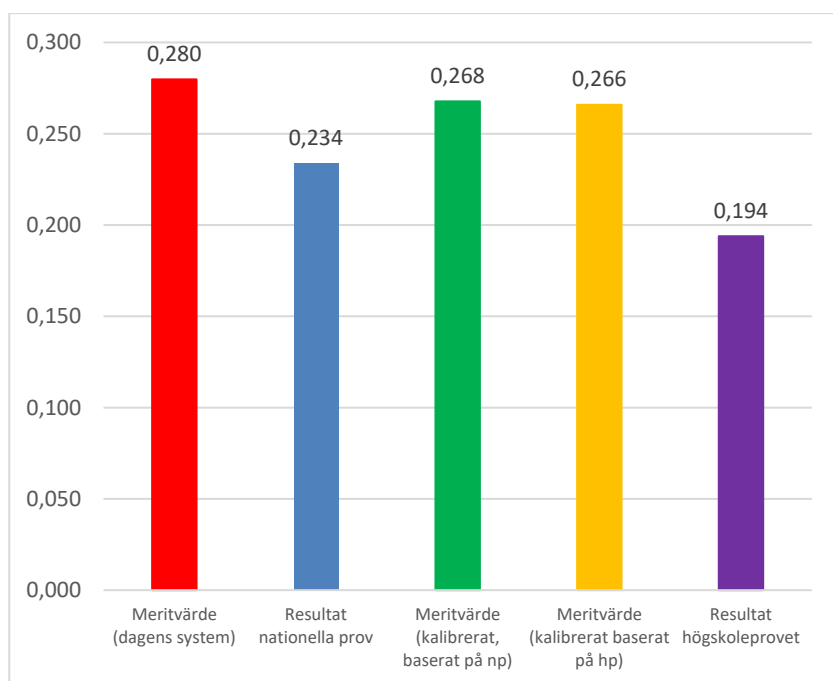
Sammantaget är bedömningen utifrån dessa data att högskoleprovet i sin nuvarande utformning inte i tillräcklig grad mäter samma kunskaper som betygen avser att mäta. Det bedöms därför som olämpligt att använda det för meritvärdeskalibrering. Det NP-kalibrerade meritvärdet faller bättre ut baserat enbart på dessa korrelationer. De framstår som rimliga eftersom de ligger inom samma intervall som de korrelationer som vi historiskt observerar mellan provbetyg och kursbetyg i olika ämnen.

⁵ Med ”sann kunskap” menas den genomsnittliga nivå som eleven skulle fått vid ett oändligt antal mätningar av ett oändligt antal betygssättande lärare. Värdet är alltså bara ett teoretiskt, latent värde.

Korrelation mellan studieresultat från gymnasieskolan och studieframgång i högskolan

I figur 1.20 visas korrelationen mellan olika mått på studieresultat från gymnasieskolan och studieframgång på högskolan. Det mått som används för studieframgång på högskolan är andel tagna högskolepoäng det första året och tar inte hänsyn till att olika kurser på högskolan kan vara olika svåra. Studieframgången utgår också från den antagning som skett med dagens betygssystem, vilket innebär att elever som kanske skulle kommit in på sin önskade utbildning med mer rättvisa betyg kanske inte gjort det. Detta kan i sin tur påverka studiemotivationen att klara kurserna på en mindre önskad utbildning. Resultaten ska därför tolkas med försiktighet.

Figur 1.20 Korrelation mellan studieresultat i gymnasieskolan och studieframgång på högskolan



Anm.: I det kalibrerade meritvärde baserat på resultat från de nationella proven ingår även elevens individuella resultat på dessa prov med vikten 30 %.

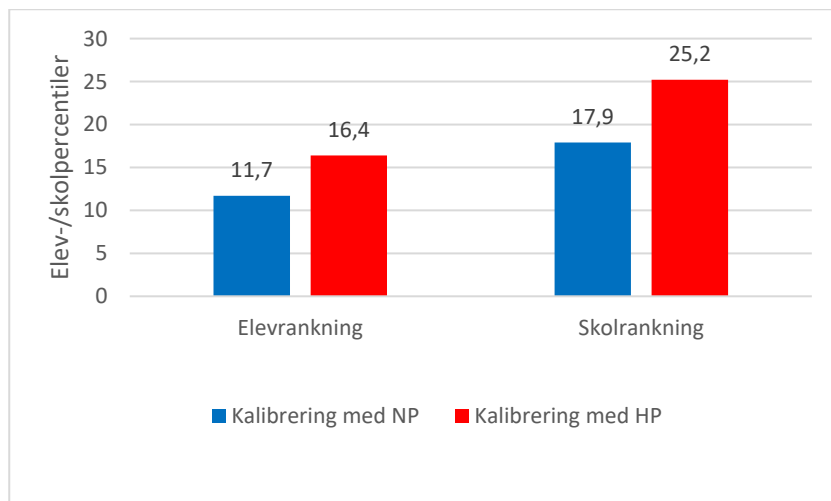
Källor: Utredningens egna beräkningar utifrån Skolverkets och Umeå universitets statistik.

Figur 1.20 visar att korrelationen mellan meritvärde, oavsett om det är kalibrerat eller ej, och studieframgång på universitetet är högre än resultaten från nationella prov samt resultatet från högskoleprovet. Det indikerar att lärarbetygen innehåller mycket värdefull information om elevers kunskaper som inte fullt ut framkommer om man enbart skulle förlita sig på nationella slutprov eller högskoleprovet. Det betyder inte nödvändigtvis att proven inte mäter rätt saker utan att det antagligen har mer att göra med det stora mätfel som uppstår när elevernas kunskaper baseras på en eller några få mätningar. Vi kan också notera att de nationella proven uppvisar en högre korrelation med studieframgång på högskolan jämfört med högskoleprovet.

Det kan också noteras att det kalibrerade meritvärdet ligger på ungefär samma nivå som meritvärdet enligt dagens system men inte högre, vilket kanske hade förväntats. Det finns antagligen flera förklaringar till det. Den meritvärdeskalibrering vi har kunnat genomföra är gjord på befintliga data, vilka är långt ifrån perfekta. Till exempel har vi inte haft tillgång till provresultaten uttryckta i råpoäng utan endast i form av provbetyg, vilket innebär att mycket information slängs bort. Andra svagheter är att proven inte är centralt rättade och kalibreringen utgår endast från nationella prov i tre ämnen – svenska, matematik och engelska – medan en implementerad meritvärdeskalibreringsmodell kommer att baseras på nationella provresultat i ytterligare tre ämnen. Det är alltså troligt att korrelationen kommer att vara något högre med meritvärdeskalibrering om den väl införs på riktigt jämfört med vad denna tentativa analys visar.

Systematiska skillnader i elev- respektive skolrankning

Figur 1.21 Genomsnittlig skillnad i ranking för elever respektive skolor i riket jämfört med originalmeritvärdet



Anm.: Staplarna anger genomsnittlig skillnad i percentilrankning av elever respektive skolor mellan dagens meritvärde och ett kalibrerat meritvärde, mätt som standardavvikelsen i dessa skillnader på både elev- och skolnivå.

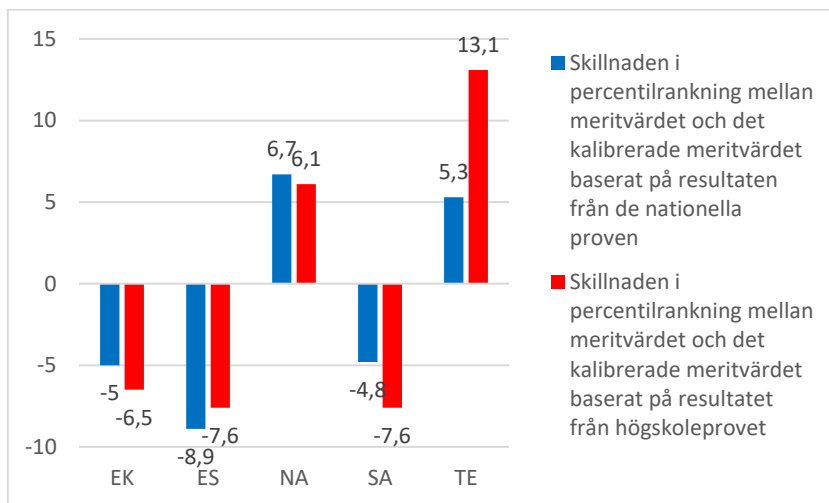
Källor: Utredningens egna beräkningar utifrån Skolverkets och Umeå universitets statistik.

Figur 1.21 visar *standardavvikelsen* i differensen i percentilrankning mellan meritvärdet och det kalibrerade meritvärdet baserat på NP respektive HP för både elevnivå och skolnivå. Eftersom rankingen är gjord för samtliga elever/skolor i urvalet (årsvis) så kommer medelvärden av differensen i percentilrankning alltid att vara 0. Standardavvikelsen i percentilrankningen ger däremot information om den genomsnittliga absoluta avvikelsen i denna rankning för en elev/skola (dvs. oavsett om förändringen är positiv eller negativ). Att införa meritvärdeskalibrering utifrån nationella slutprov beräknas således innebära en förändring i elevers rankning, med i genomsnitt ungefär 12 elevpercentiler. För skolor skulle reformen innebära en förändring med i genomsnitt 18 skolpercentiler. Förändringarna får anses vara betydande. Motsvarande genomsnittliga skillnader med det högskoleprovskalibrerade meritvärdet är ännu större, 16 elev- respektive 25 skolpercentiler. Detta är på inget sätt överraskande eftersom vi i figur 1.19 kunde se att korrelationen med meritvärdet var betydligt lägre för det HP-kalibrerade meritvärdet.

Systematiska skillnader per gymnasieprogram

När analysen görs programvis summerar medelvärdet inte nödvändigtvis till 0. Därför redovisas i figur 1.22 och framåt medelvärden för respektive undergrupp, i detta fall gymnasieprogram.

Figur 1.22 Systematisk skillnad i elevrankning per gymnasieprogram och kalibreringsmodell

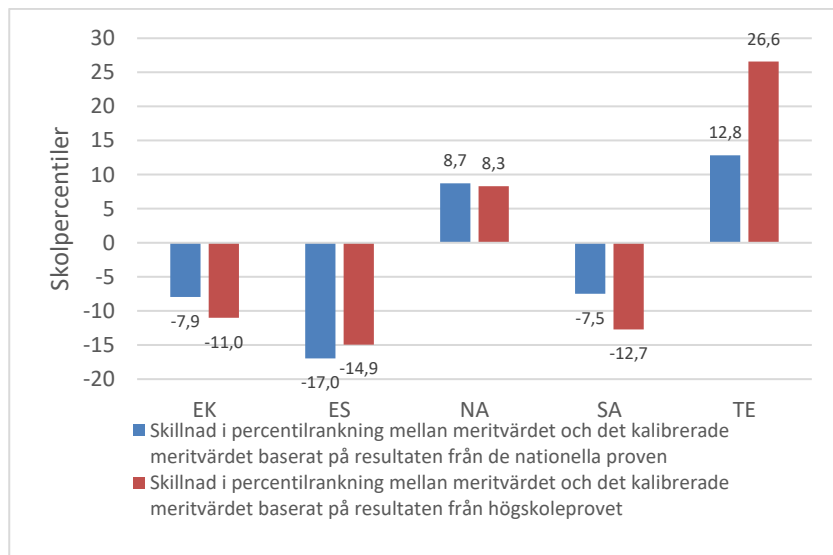


Anm.: Staplarna anger systematiska skillnader i elevrankning respektive skolrankning mellan dagens meritvärde och ett kalibrerat meritvärde, dels kalibrerat efter resultaten på nationella prov, dels kalibrerat på resultaten från högskoleprovet.

Källor: Utredningens egna beräkningar utifrån Skolverkets och Umeå universitets statistik.

Båda figurer, 1.22 och 1.23, visar samma generella mönster och mönstret är på det hela taget detsamma för båda varianterna av meritvärdeskalibrering. Elever på NA- respektive TE-programmen beräknas i genomsnitt tjäna på meritvärdeskalibreringen eftersom deras ranking i genomsnitt kommer att förbättras med ungefär sju respektive 5–13 elevpercentiler. Elever på övriga program beräknas däremot att i genomsnitt förlora. Skolrankningen påverkas ännu mer. Skolor med teknikprogram väntas klättra i genomsnitt 13–27 skolpercentiler medan ES-program förlorar nästan lika mycket i ranking. Observera att när vi talar om skolrankning och skolpercentiler menar vi att samtliga elever på samma program på samma skola utgör en programskola i analysen och denna grupp utgör basen för meritvärdeskalibreringen på gruppnivå.

Figur 1.23 Systematisk skillnad i skolrankning per program och kalibreringsmodell



Anm.: Staplarna anger systematiska skillnader i skolrankning respektive skolrankning mellan dagens meritvärde och ett kalibrerat meritvärde utifrån de nationella proven respektive högskoleproven.

Källor: Utredningens egna beräkningar utifrån Skolverkets och Umeå universitets statistik.

Resultaten reflekterar det faktum att betygssättningen i dag är relativt strängare på NA- och TE-programmet, givet samma resultat på de gemensamma proven. Genom att vi har placerat de högre matematikkursprovsresultaten på samma skala (som matematik 1-resultaten) förstärker vi dock effekten av meritvärdeskalibreringen. I nuvarande system "straffas" elever som läser en svårare kurs genom att de kan få svårare att få ett högt betyg. I meritvärdeskalibreringsmodellen har vi med hjälp av det gemensamma matematik 1-provet (vi har slagit ihop matematik 1a–c och antagit att de är samma prov) kunnat projicera de högre matematikprovresultaten på samma skala som matematik 1-provet. Eftersom NA- och TE-eleverna i genomsnitt presterar bättre än övriga program på matematik 1-provet, kommer deras resultat på t.ex. matematik 4-provet att skalas upp till en genomsnittlig nivå som motsvarar samma grupps genomsnittliga resultat på matematik 1-provet.

Dessutom neutraliserar vi den negativa betygseffekten av en relativt svårare kurs genom att vi årsvis standardiserar samtliga prov. Detta innebär att elever som läser till exempel fysik, där det är svårare

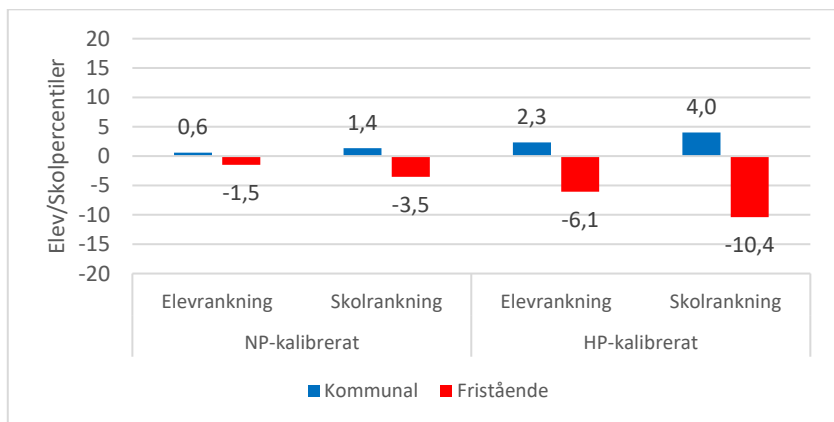
att få ett A än i t.ex. samhällskunskap, tack vare standardiseringen, kommer att ha lika högt genomsnittligt provresultat på nationell nivå som eleverna som läser samhällskunskap.

1.3.5 Konsekvenser för olika undergrupper av elever

Systematiska skillnader efter huvudman

Elever i fristående skolor beräknas förlora i ranking medan rankingen för elever i kommunala skolor förbättras (figur 1.24). Staplarna anger systematiska (genomsnittliga) skillnader i elevrankning respektive skolrankning mellan dagens meritvärde och ett kalibrerat meritvärde. Skillnaden slår igenom något mer för NA- respektive TE-elever men något mindre för EK-elever (figur 1.25).

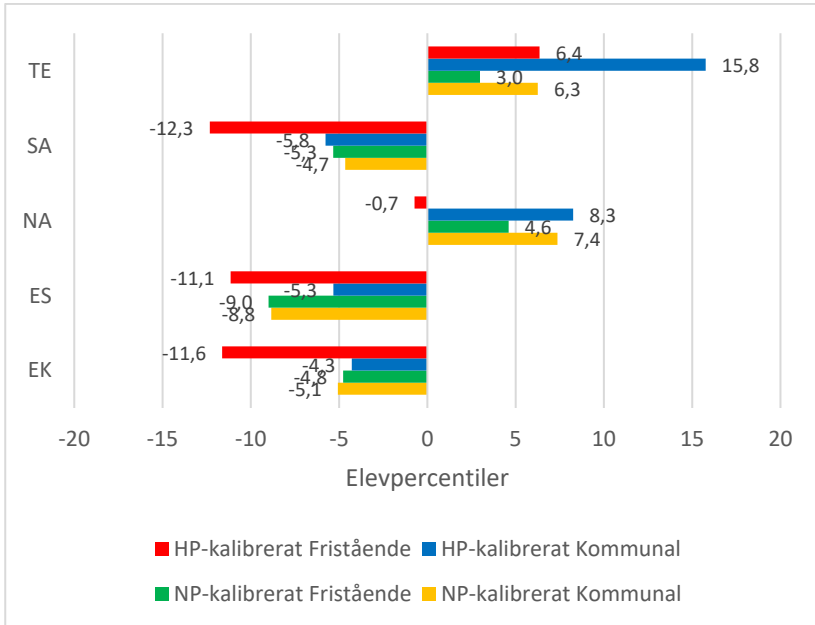
Figur 1.24 Genomsnittlig skillnad i elevrankning och skolrankning efter huvudman och kalibreringsmodell jämfört med nuvarande meritvärde



Anm.: Staplarna anger systematiska (genomsnittliga) skillnader i elevrankning respektive skolrankning mellan dagens meritvärde och ett kalibrerat meritvärde, dels baserat på resultaten från de nationella proven (NP), dels baserat på resultatet från högskoleprovet (HP).

Källor: Utredningens egna beräkningar utifrån Skolverkets och Umeå universitets statistik.

Figur 1.25 Systematiska skillnader i elevrankning efter huvudman och program



Källor: Utredningens egna beräkningar utifrån Skolverkets och Umeå universitets statistik.

Vad gäller kalibreringsmodell blir skillnaderna mellan fristående och kommunala skolor betydligt mer framträdande med HP-kalibrering. På programnivå syns tydliga skillnader mellan kommunala och fristående skolor för samtliga program. Det är inte på något sätt självklart hur denna diskrepans ska förstås, men en delförklaring skulle kunna vara att proven i denna analys bygger på provresultat som inte är centralt rättade utan oftast bedömda av lärare på elevens egen skola. Tidigare studier har visat att lärare på fristående skolor tenderar att vara mer generösa i provbedömningen.⁶ Högskoleprovet är däremot helt oberoende av betygssättande lärare och då skulle en del av de större skillnaderna kunna bero på detta. I ett framtida betygssystem föreslås alla nationella slutprov att vara centralt rättade så detta resultat, om det stämmer, bör inte ses som ett argument för en högskoleprovsbaserad kalibreringsmodell.

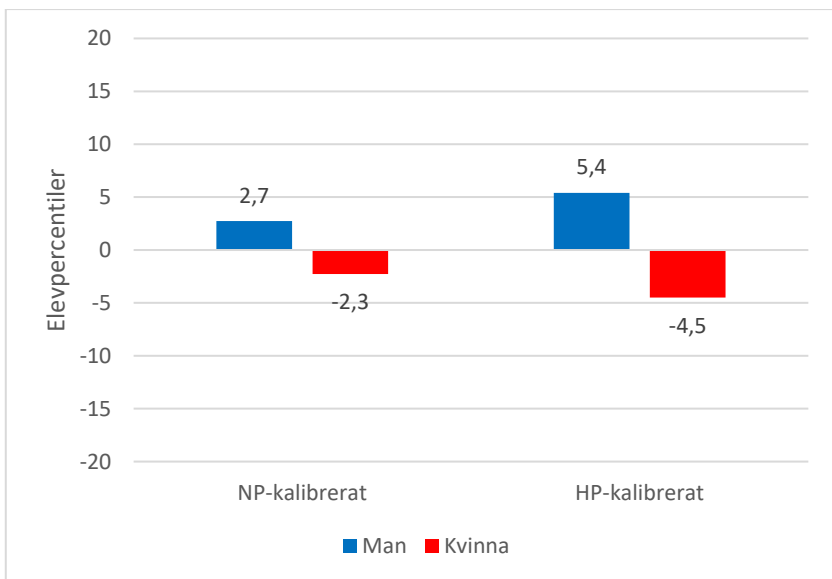
⁶ Tyrefors Hinnerich & Vlachos (2016).

Systematiska skillnader efter kön

Totalt framgår att reformeffekten blir två och en halv elevpercentiler till pojkarnas fördel i den NP-baserade kalibreringsmodellen och fem elevpercentiler i den HP-baserade kalibreringsmodellen. När skillnaderna mellan könen delas upp per program minskar skillnaderna något, oavsett kalibreringsmodell.

Dessa resultat kan framstå som en betydande försämring för flickor, men eftersom flickors genomsnittliga percentilrankning ligger cirka 10 elevpercentiler högre än pojkars i dagens betygssystem (54,6 respektive 44,5) kommer flickors rankning fortfarande att vara högre än pojkars på gruppnivå med en NP-baserad kalibreringsmodell. Flickornas fördel består men halveras till cirka fem elevpercentiler (52,4 respektive 47,1) för den NP-baserade kalibreringsmodellen. Med den HP-baserade kalibreringsmodellen kommer flickornas fördel i percentilrankningen att elimineras. Observera att dessa jämförelser inte framgår i figurerna.

Figur 1.26 Systematisk skillnad i elevrankning efter elevens kön

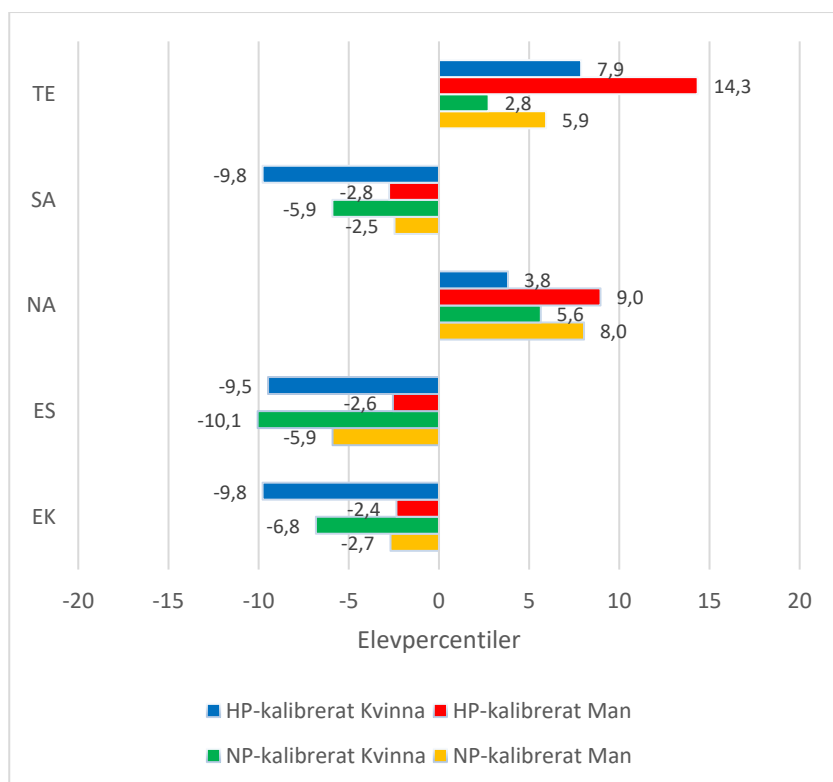


Anm.: Staplarna anger systematiska skillnader i elevrankning respektive skolrankning mellan dagens meritvärde och ett kalibrerat meritvärde utifrån medelvärdet.

Källor: Utredningens egna beräkningar utifrån Skolverkets och Umeå universitets statistik.

Observera att vi egentligen inte kan veta, utifrån dessa analyser, om de potentiella förändringarna innebär en förbättrad eller försämrad likvärdighet utifrån ett könsperspektiv. En möjlig tolkning är att proven gynnar män, en annan att lärarsatta betyg gynnar kvinnor.

Figur 1.27 Systematiska skillnader i elevrankning efter elevens kön och program



Anm.: Staplarna anger systematiska skillnader i elevrankning respektive skolrankning mellan dagens meritvärde och ett kalibrerat meritvärde utifrån högskoleproven respektive de nationella proven.

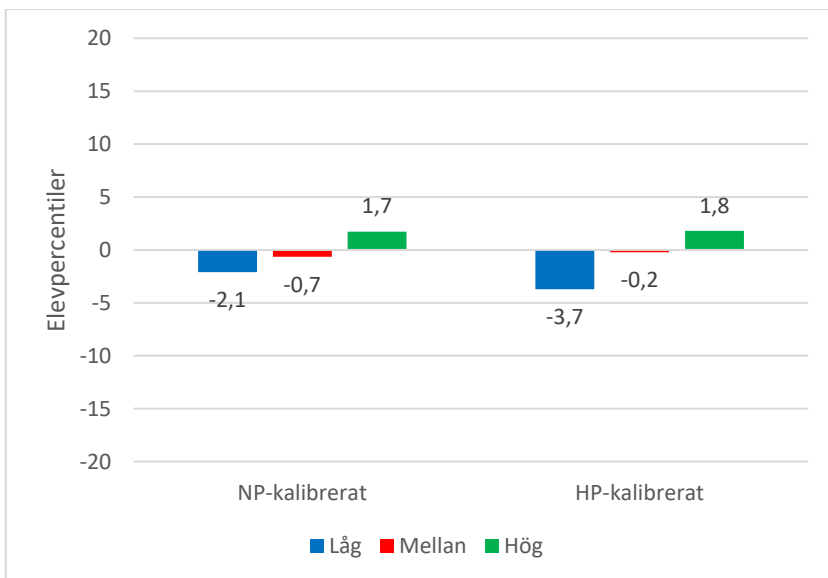
Källor: Utredningens egna beräkningar utifrån Skolverkets statistik och Umeå universitets statistik.

Systematiska skillnader efter föräldrars utbildningsnivå

I figur 1.28 visas att elever vars föräldrar har en högre utbildningsnivå i genomsnitt beräknas förbättra sin rankning med reformen (1,7 percentiler med nationella prov och 1,8 percentiler med högskoleprov) och det omvända väntas ske för elever vars föräldrar har en lägre utbildningsnivå (2,1 percentiler med nationella prov och 3,7 percentiler

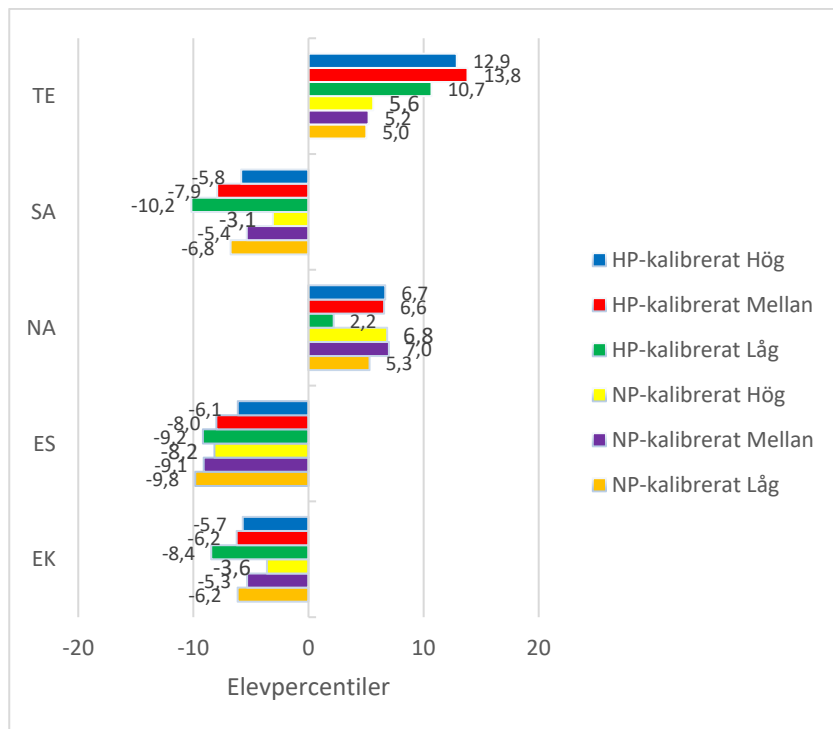
med högskoleprov). Effekten av reformen blir mindre när man analyserar motsvarande skillnader inom respektive program. Den rimligaste förklaringen till detta mönster är att det finns ett betydande inslag av relativ betygssättning mellan skolor i dagens system. Med meritvärdeskalibreringen elimineras denna olikvärdighet. Föräldrars utbildningsnivå samvarierar (positivt) med elevers prestationsnivå och elever vars föräldrar har förhållandevis låg utbildningsnivå går i större utsträckning på lågpresterande skolor (skolor vars genomsnittliga prestationsnivå är låg). Därmed kommer elever vars föräldrar har låg utbildningsnivå i genomsnitt att förlora i elevrankning. Det kan låta negativt men den relativa betygssättningen elimineras och likvärdigheten i betygssättningen ökar.

Figur 1.28 Systematiska skillnader i elevrankning efter föräldrars utbildningsnivå



Källor: Utredningens egna beräkningar utifrån Skolverkets och Umeå universitets statistik.

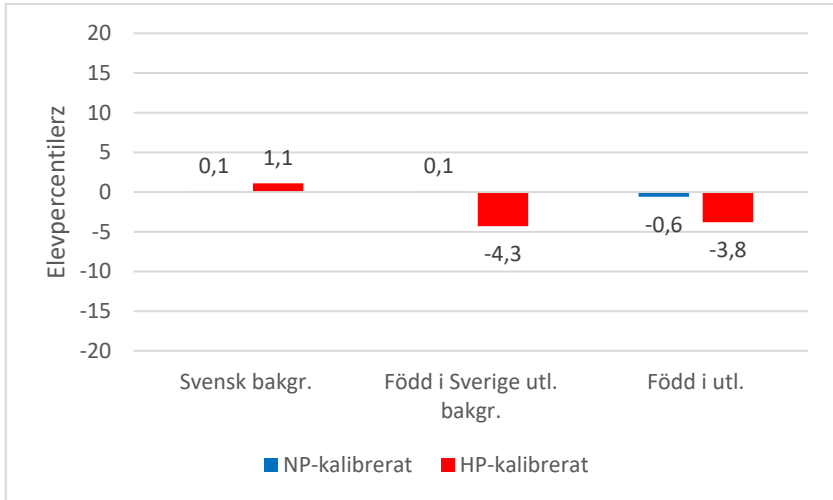
Figur 1.29 Systematiska skillnader i elevrankning efter föräldrars utbildningsnivå och program



Källor: Utredningens egna beräkningar utifrån Skolverkets och Umeå universitets statistik.

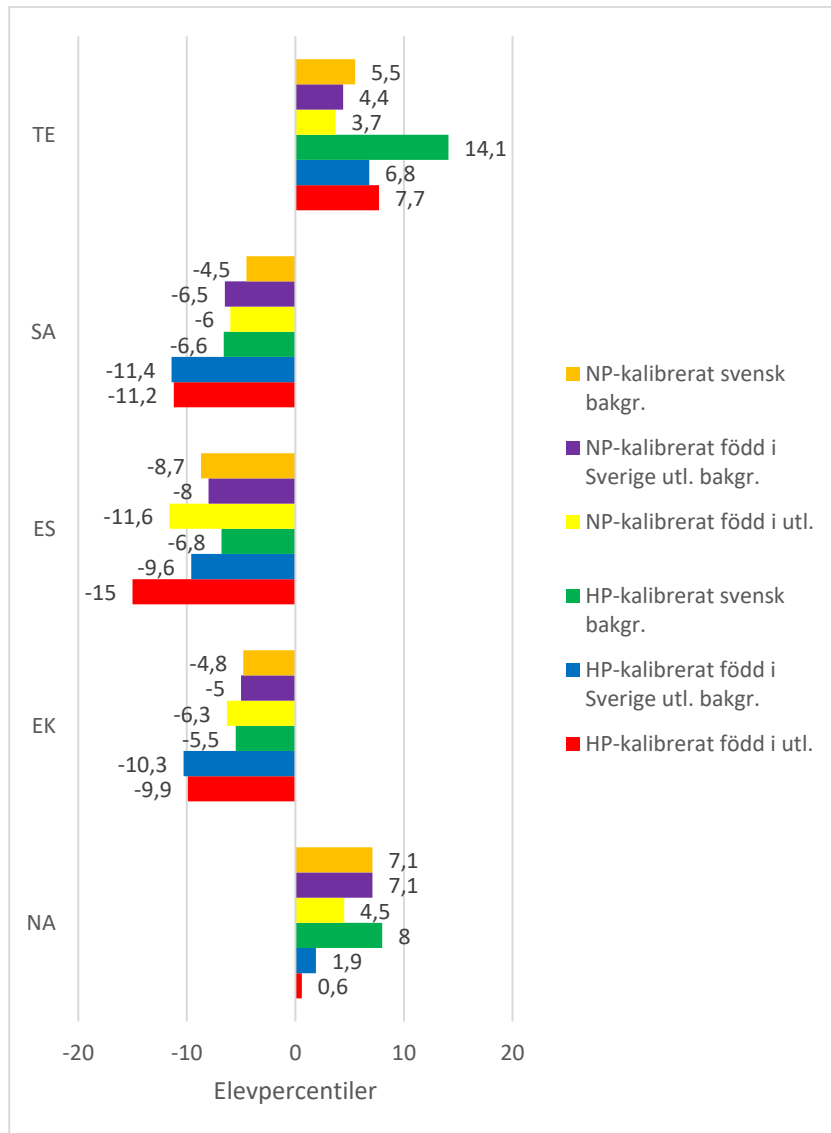
Systematiska skillnader i elevrankning efter migrationsbakgrund

Figureorna 1.30 och 1.31 visar att den totala effekten för elever med olika migrationsbakgrund är minimal med den NP-baserade kalibreringsmodellen men något större med den HP-baserade kalibreringsmodellen där såväl utlandsfödda elever som elever födda i Sverige med utländsk bakgrund tappar ungefär fyra elevpercentiler. När reformeffekterna analyseras efter program syns något större skillnader, men de måste ändå anses som relativt små när det gäller den NP-baserade kalibreringsmodellen men större för den HP-baserade kalibreringsmodellen.

Figur 1.30 Systematiska skillnader i elevrankning efter migrationsbakgrund

Källor: Utredningens egna beräkningar utifrån Skolverkets och Umeå universitets statistik.

Figur 1.31 Systematiska skillnader i elevrankning efter migrationsbakgrund och program



Källor: Utredningens egna beräkningar utifrån Skolverkets statistik och data från Umeå universitet.

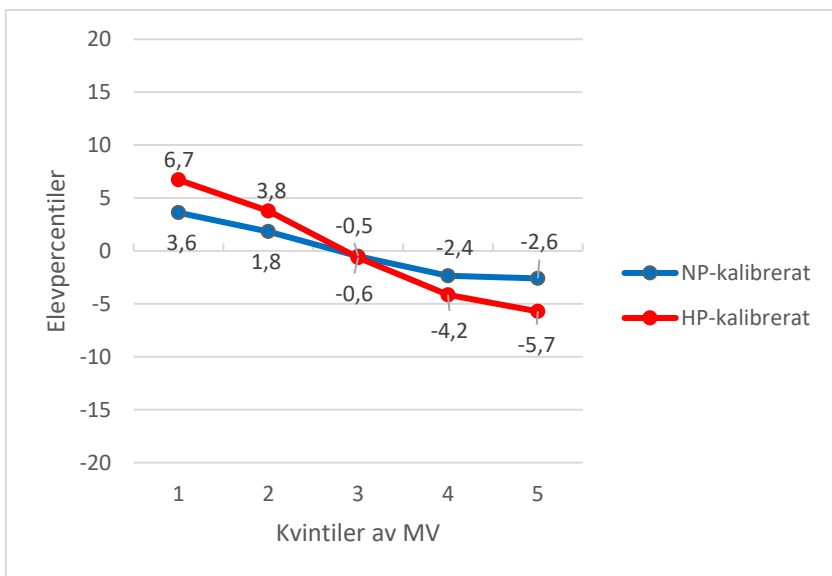
Systematiska skillnader i elevrankning efter prestationsnivå

Vi har undersökt hur elever som presterar på olika nivåer påverkas av reformen genom att dela upp eleverna i kvintiler, dvs. fem lika stora grupper rangordnade efter deras meritvärde i figurerna 1.32–1.34. Den femtedel av eleverna som har lägst meritvärde tillhör kvintil 1 medan den femtedel av eleverna som har högst meritvärde tillhör kvintil 5. Lågpresterande elever förbättrar i genomsnitt sin elevrankning medan högpresterande elever (utifrån gällande meritvärde) i genomsnitt försämrar sin elevrankning.

När analysen görs uppdelat på program är mönstret på det stora hela det samma förutom att högpresterande elever på NA- och TE-programmen inte förlorar i elevrankning.

Mönstret från tidigare figurer (1.21–1.31) står sig: den HP-baserade kalibreringsmodellen ger ett större utslag jämfört med den NP-baserade kalibreringsmodellen.

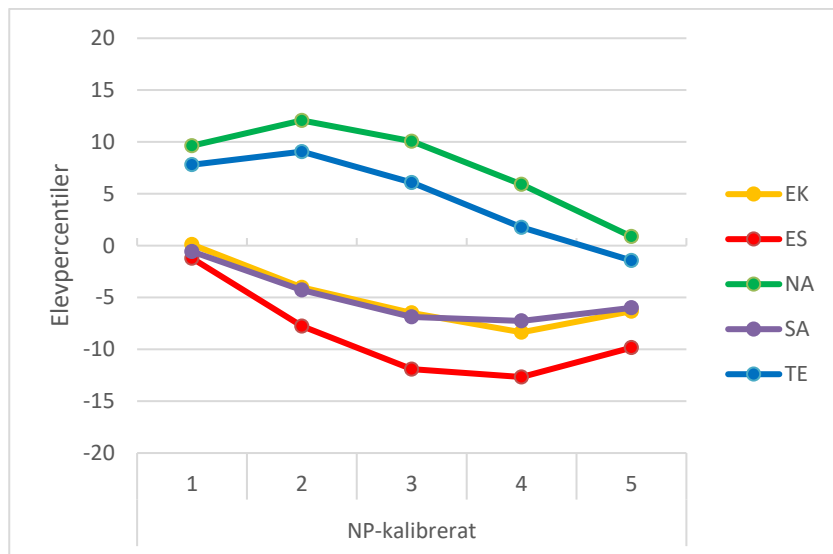
Figur 1.32 Systematiska skillnader i elevrankning efter kvintiler av elevers meritvärde



Anm.: Rangordning efter kvintiler innebär att eleverna delas upp i fem lika stora grupper (20 % i varje kvintil) baserat på deras meritvärde. Kvintil 1 innefattar de 20 % av eleverna som har lägst meritvärde och kvintil 5 de 20 % som har högst meritvärde.

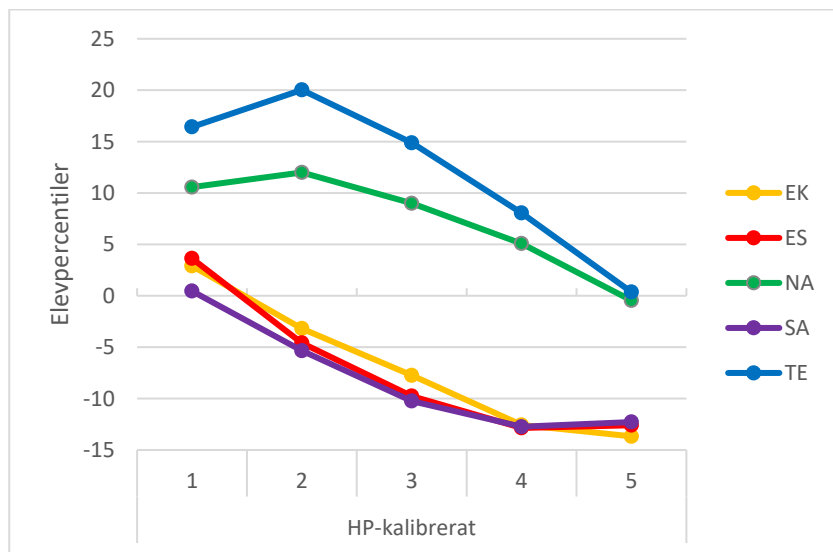
Källor: Utredningens egna beräkningar utifrån Skolverkets och Umeå universitets statistik.

Figur 1.33 Systematiska skillnader i elevrankning efter kvintiler av elevers meritvärde och program, baserat på NP-kalibrering



Källor: Utredningens egna beräkningar utifrån Skolverkets och Umeå universitets statistik.

Figur 1.34 Systematiska skillnader i elevrankning efter kvintiler av elevers meritvärde och program, baserat på HP-kalibrering



Källor: Utredningens egna beräkningar utifrån Skolverkets och Umeå universitets statistik.

1.3.6 Regressionsanalys – systematiska effekter

Vi kan nu sammanställa vilka faktorer som får störst effekt vid ett införande av meritvärdeskalibrering. I tabellerna 1.6 och 1.7 visas effekterna av meritvärdeskalibrering på olika elevgruppers percentilrankning när hänsyn tas till övriga faktorer, dvs. de övriga variablerna som redovisas. I tabell 1.6 är meritvärdeskalibreringen baserad på resultaten från de nationella proven medan den i tabell 1.7 är baserad på resultaten från högskoleprovet. Meritvärdeskalibrering påverkar framför allt elever på natur- och teknikprogrammet positivt. De som förlorar mest, jämfört med dagens meritvärden (jämförelsetal), är elever på fristående skolor samt elever med de högsta betygen på skolor som sätter höga betyg i relation till elevers prestation på nationella prov och högskoleprov.

Tabell 1.6 Resultat från regressionsanalys baserad på NP-kalibrering

Utfallsvariabel = skillnad i percentilrankning mellan okalibrerat meritvärde (= jämförelsetal i dagens system) och kalibrerat meritvärde där även elevens individuella resultat från de nationella proven vägs in med 30 % (elevpercentiler)

	Koefficient	Stand. koefficient	t-värde	p-värde
Konstant	-19,54		-46,9	0,00
ES	-2,41	-0,045	-8,7	0,00
SA	0,80	0,03	5,0	0,00
NA	7,32	0,29	41,3	0,00
TE	7,74	0,24	41,0	0,00
Kvinna	-3,37	-0,14	-28,6	0,00
Utlandsfödd	0,20	0,01	1,1	0,28
Född i Sverige med utlandsfödda föräldrar	1,12	0,03	6,2	0,00
Föräldrars genomsnittliga utbildningsnivå	-0,17	-0,02	-3,4	0,00
Fristående huvudman	-2,59	-0,10	-20,4	0,00
Elevens NP-resultat	0,11	0,14	22,8	0,00
Skolans NP-resultat	0,28	0,21	29,9	0,00
Eleven med högst okalibrerat meritvärde på skolan	-4,91	-0,11	-22,1	0,00

Anm.: $R^2 = 0,34$. Den standardiserade koefficienten är effekten uttryckt som antal standard avvikelser i utfallsvariabeln (y i en regression) när den oberoende variabeln (x-variabeln i en regression) ökar med en standardavvikelse. Ekonomiprogrammet utgör referensgrupp och syns därför inte i tabellen.

Tabell 1.7 Resultat från regressionsanalys baserad på HP-kalibrering

Utfallsvariabel = skillnad i percentilrankning mellan okalibrerat meritvärde och kalibrerat meritvärde där även elevens individuella resultat vägts in med 30 % (elevpercentiler)

	Koefficient	Stand. koefficient	t-värde	p-värde
Konstant	-25,53		-42,8	0,00
ES	0,57	0,01	1,6	0,11
SA	0,23	0,01	1,1	0,26
NA	5,79	0,16	24,4	0,00
TE	12,96	0,28	50,3	0,00
Kvinna	-5,10	-0,16	-33,1	0,00
Utlandsfödd	-0,73	-0,01	-2,9	0,00
Född i Sverige med utlandsfödda föräldrar	-0,72	-0,01	-2,9	0,00
Föräldrars genomsnittliga utbildningsnivå	-0,75	-0,06	-11,5	0,00
Fristående huvudman	-7,55	-0,21	-47,1	0,00
Elevens NP-resultat	0,32	0,29	51,4	0,00
Skolans NP-resultat	0,29	0,15	22,7	0,00
Eleven med högst okalibrerat meritvärde på skolan	-7,62	-0,12	-27,2	0,00

Anm.: $R^2 = 0,44$. Den standardiserade koefficienten är effekten uttryckt som antal standardavvikelser i utfallsvariabeln (y i en regression) när den oberoende variabeln (x-variabeln i en regression) ökar med en standardavvikelse. Ekonomiprogrammet utgör referensgrupp och syns därför inte i tabellen.

Den avgörande skillnaden mellan resultaten från regressionsanalyserna jämfört med resultaten som framgår av analyserna som redovisas i tidigare figurer är att vi här tar hänsyn till övriga variabler när effekten av en viss variabel skattas. Regressionskoefficienterna är därmed att betrakta som ett slags nettoeffekter, givet att modellen är korrekt specificerad.

Oavsett kalibreringsmodell (NP- eller HP-baserad) är de generella mönstren desamma och går i stort sett i linje med de resultat vi såg i de föregående analyserna där ingen hänsyn tagits till övriga variabler. När det gäller programindikatorerna ska de tolkas som skillnad jämfört med EK-programmet som är referensgrupp. NA- och TE-programmen påverkas positivt jämfört med EK-programmet. En skillnad är att det inte framgår några negativa effekter av att vara utlandsfödd eller född i Sverige med utländsk bakgrund med den NP-kalibrerade modellen. Med HP-kalibrering är däremot effek-

ten negativ för båda grupperna. För kvinnor är den negativa effekten större med HP-kalibrering än med NP-kalibrering, men skillnaden är mindre jämfört med ”bruttoeffekterna” i tidigare figurer, speciellt för den HP-kalibrerade modellen.

En viktig skillnad är att föräldrarnas utbildningsnivå har en negativ effekt, vilket inte var fallet i de enkla sambandsanalyserna (figur 1.28 och 1.29). Effekten är dock relativt svag (se den standardiserade effekten då detta är en kontinuerlig variabel och enheterna är svårtolkade). Att gå i en fristående skola har en negativ effekt på 2,6 respektive 7,5 elevpercentiler beroende på modell.

Den kanske mest intressanta effekten gäller hur rankningen påverkas av skolans genomsnittliga resultat på slutproven med en standardiserad effektstorlek på 0,21, vilket vi bedömer som en ganska stor effektstorlek. Observera att detta inte endast ska tolkas som att ju bättre skolan skriver på proven, desto högre rankning för eleverna eftersom utfallsvariabeln är *förändring* i elevers percentilrankning jämfört med percentilrankning baserad på det ursprungliga meritvärdet. Denna relativt stora effekt speglar det starka inslaget av relativ betygssättning i nuvarande betygssystem. Skolor som har en hög generell prestationsnivå är i genomsnitt restriktiva i betygssättningen och vice versa. När meritvärdena kalibreras efter hur skolan presterar på slutproven elimineras detta inslag, därav den stora positiva effekten. Det finns även en positiv effekt på elevnivå för elever som, allt annat lika, skriver bra på de nationella slutproven. Det ska alltså tolkas som att elever som skriver bra på de nationella slutproven får utdelning för detta.

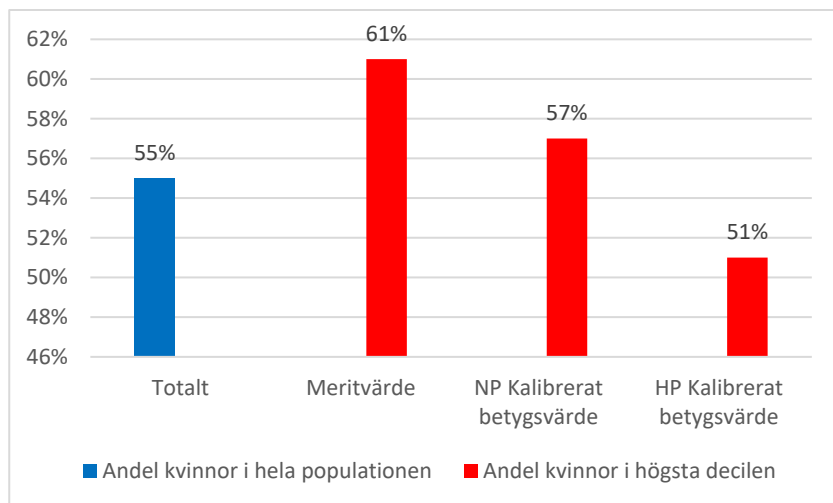
Slutligen kan vi konstatera en negativ effekt för elever som har det högsta meritvärdet på skolan. Det skulle kunna bero på takeffekter som kan uppstå när betygen inte lyckas differentiera en högpresterande elev från övriga elever på skolan om lärarna har en relativt sett generös betygssättning i förhållande till hur eleverna presterar på proven. Effekten är större med den HP-kalibrerade modellen.

1.3.7 Elevsammansättning för den högsta decilen av eleverna utifrån meritvärde och kalibrerat meritvärde

I detta avsnitt redovisas hur elevsammansättningen i den högsta decilen (de 10 % högst presterande eleverna) kan komma att förändras med en meritvärdeskalibreringsmodell. Den blå stapeln i respek-

tive figur anger referensandelen, dvs. andelen i hela populationen som ingår i vårt urval och som motsvarar alla elever på de fem största högskoleförberedande programmen.

Figur 1.35 Andel kvinnor i den högsta decilen beroende på betygsmått

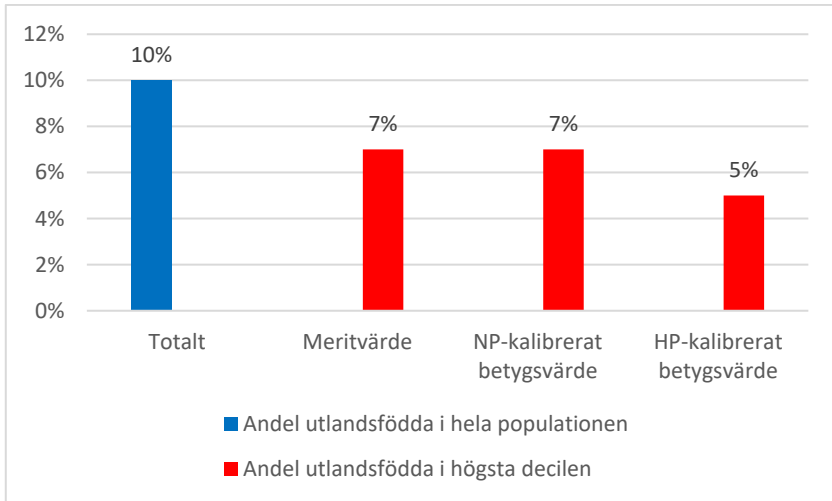


Anm.: Den blå stapeln anger andelen kvinnor i hela populationen elever på de fem vanligaste högskoleförberedande programmen. De röda staplarna anger andelen kvinnor i den högsta decilen beroende på betygsmått. Meritvärde = ursprungligt meritvärde (okalibrerat) NP-kalibrerat betygsvärde = Kalibrerat betygsvärde baserat på nationella prov och där elevens individuella resultat på dessa vägs in med 30 %. HP-kalibrerat betygsvärde = Kalibrerat betygsvärde baserat på högskoleprovet och där elevens individuella resultat på detta vägs in med 30 %.

Källor: Utredningens egna beräkningar utifrån Skolverkets och Umeå universitets statistik.

Figur 1.35 visar att andelen kvinnor på de högskoleförberedande programmen totalt sett, är något överrepresenterade och uppgår till 55 %. I den högsta decilen är kvinnor ännu mer överrepresenterade, baserat på dagens meritvärde och uppgår till 61 %. Med meritvärdeskalibrering skulle den andelen minska till 57 % enligt våra beräkningar. Med en högskoleprovsbaserad kalibreringsmodell skulle andelen minska till 51 %, vilket innebär att kvinnor i den högsta decilen i praktiken skulle vara underrepresenterade eftersom de utgör 55 % av alla elever på högskoleförberedande program.

Figur 1.36 Andel utlandsfödda elever i den högsta decilen beroende på betygsmått

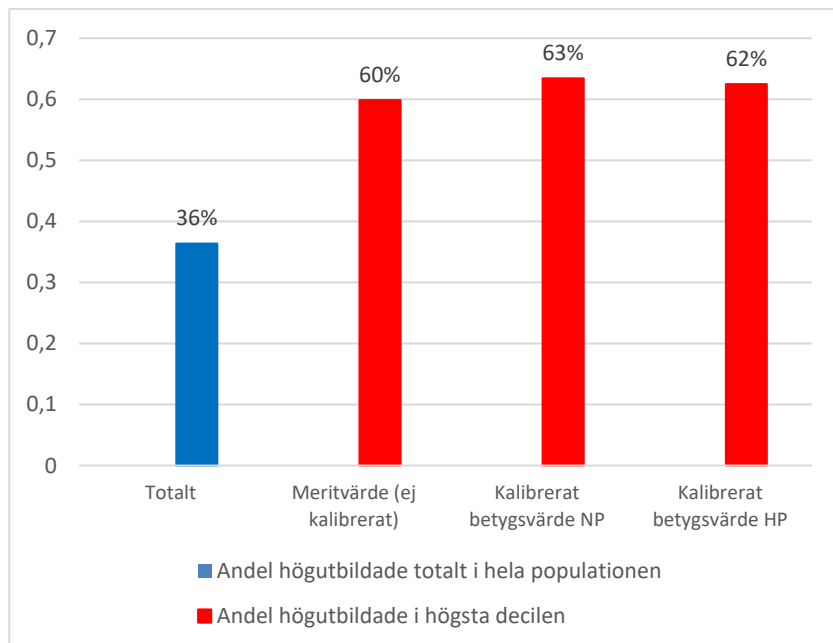


Anm.: Den blå stapeln anger andelen utlandsfödda i hela populationen elever på de fem vanligaste högskoleförberedande programmen. De röda staplarna anger andelen utlandsfödda i den högsta decilen beroende på betygsmått. Meritvärde = Ursprungligt meritvärde (okalibrerat). NP-kalibrerat betygsvärde = Kalibrerat betygsvärde baserat på nationella prov och där elevens individuella resultat på dessa vägts in med 30 %. HP-kalibrerat betygsvärde = Kalibrerat betygsvärde baserat på högskoleprovet och där elevens individuella resultat på detta vägts in med 30 %.

Källor: Utredningens egna beräkningar utifrån Skolverkets och Umeå universitets statistik.

I figur 1.36 visas på motsvarande sätt andelen utlandsfödda i den högsta decilen. De utgör 10 % av alla elever på högskoleförberedande programmen men är underrepresenterade i den högsta decilen där de utgör 7 %. Med meritvärdeskalibrering utifrån nationella prov skulle den andelen inte ändras men med en högskoleprovsbaserad kalibrering beräknas andelen utlandsfödda i den högsta decilen minska till 5 %.

Figur 1.37 Andel elever med högutbildade föräldrar i den högsta decilen beroende på betygsmått

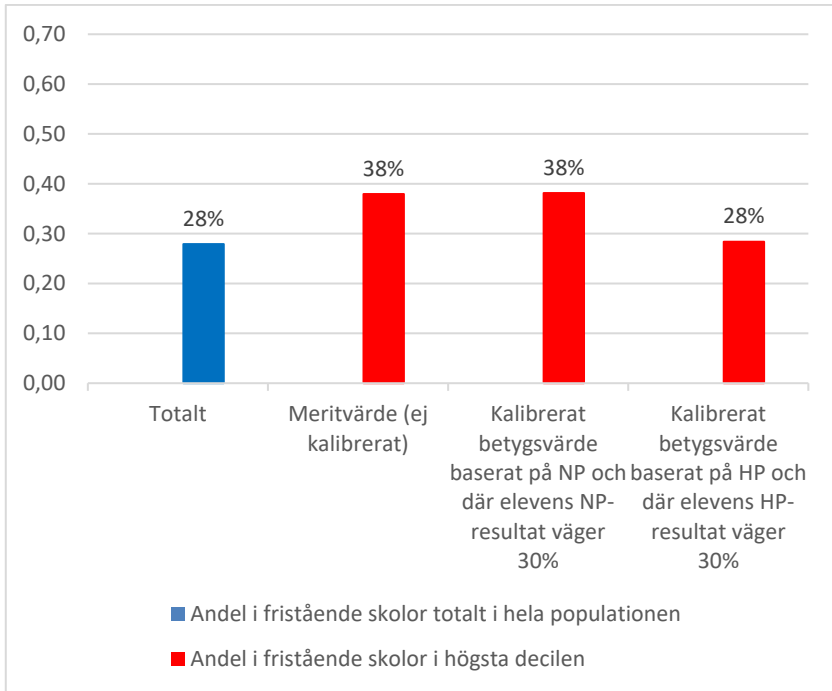


Anm.: Den blå stapeln anger andelen högutbildade i hela populationen elever på de fem vanligaste högskoleförberedande programmen. De röda staplarna anger andelen högutbildade i den högsta decilen beroende på betygsmått. Kalibrerat betygsvärde NP innebär ett betygsvärde baserat på kalibrering med nationella provresultat och där elevens individuella resultat på dessa vägs in med 30 %. Kalibrerat betygsvärde HP innebär ett betygsvärde baserat på kalibrering med högskoleprovresultat och där elevens individuella resultat på detta vägs in med 30 %.

Källor: Utredningens egna beräkningar utifrån Skolverkets och Umeå universitets statistik.

I figur 1.37 visas andelen elever med högutbildade föräldrar i högskoleförberedande program. Andelen är 36 %. I den översta decilen är dock andelen 60 % baserat på dagens meritvärde. Med meritvärdeskalibrering baserat på nationella prov skulle andelen öka ytterligare något till 63 % och till 62 % med högskoleprovsbaserad kalibrering.

Figur 1.38 Andel elever i fristående skolor i den högsta decilen beroende på betygsmått



Anm.: Den blå stapeln anger andelen elever i fristående skolor i hela populationen elever på de fem vanligaste högskoleförberedande programmen. De röda staplarna anger andelen elever i fristående skolor i den högsta decilen beroende på betygsmått.

Källor: Utredningens egna beräkningar utifrån Skolverkets och Umeå universitets statistik.

Slutligen visas i figur 1.38 andelen elever i fristående skolor. 28 % av alla elever på högskoleförberedande program går i fristående skolor, men andelen i den högsta decilen uppgår till 38 % baserat på dagens meritvärde. Den andelen väntas inte ändras med meritvärdeskalibrering baserat på nationella prov men beräknas sjunka till 28 % med en högskoleprovsbaserad kalibreringsmodell.

Detta sista resultat väcker en del frågor. Det är i sig inte konstigt att mönstret i den högsta decilen inte återspeglar hur effekterna av meritvärdeskalibrering slår systematiskt för hela elevkohorten. Det finns många högpresterande fristående skolor så att meritvärdeskalibreringen baserat på nationella prov inte ändrar andelen elever i den högsta decilen från fristående skolor är fullt möjligt. Det som är uppseendeväckande är att andelen sjunker så kraftigt med en hög-

skoleprovsbaserad kalibrering. Vi bedömer att ett sådant mönster skulle kunna uppstå om fristående skolor rättar sina nationella prov så pass mycket mer generöst att meritvärdeskalibreringen baserad på nationella prov inte lyckas komma åt skillnader i betygssättningen mellan kommunala och fristående skolor. Med högskoleprovet som rättas centralt får kalibreringen en helt annan effekt. Även om denna hypotes inte kan bevisas i strikt mening så är mönstret konsistent med vad vi skulle finna om hypotesen är riktig. Under alla omständigheter pekar det på vikten av att samtliga nationella prov bedöms centralt.

Sammantaget är de beräknade effekterna på elevsammansättningen i den översta decilen inte särskilt stora. Eftersom de nationella provresultat som kalibreringen baseras på endast utgörs av matematik, svenska och engelska i det befintliga nationella provsystemet (alltså saknas provresultat i historia, religion, samhällskunskap, fysik, kemi och biologi), kan effekterna på elevsammansättningen i vissa fall vara överskattade.

1.3.8 Konsekvenser för antagning till en specifik och attraktiv högskoleutbildning

I denna analys har vi utgått från ett realistiskt exempel med antagning till Handelshögskolan i Stockholm (HHS). För åren 2018 och 2019, vilka är de år som vi har maximalt med dataunderlag för, var antagningspoängen till HHS 19,17 för båda åren. Analysen syftar till att undersöka hur stor andel av eleverna som avslutade gymnasieskolan det året som skulle ha kommit in på HHS med denna antagningspoäng utifrån dagens meritvärde (jämförelsetal) men som inte skulle ha kommit in med ett kalibrerat meritvärde.⁷ Genom att skatta denna andel får vi ett mått på konsekvensen av antagning baserad på ett kalibrerat meritvärde.

Observera att vi inte i våra data har möjlighet att identifiera hur många och vilka som i slutändan antogs till och påbörjade studier vid HHS. Analysen begränsar sig till den grupp som uppfyllde antagningskravet för att komma in på utbildningen. Notera också att ana-

⁷ Vi har däremot inte tagit hänsyn till om eleverna uppfyller den särskilda behörigheten för att studera på HHS, dvs. att eleverna måste ha läst minst matematik 3 i gymnasieskolan. Analysen ska därmed ses som en mer generell kontrafaktisk analys som ger en inblick i hur relativt högpresterande elever påverkas även vad gäller intagning till andra utbildningar med höga antagningspoäng, oavsett vad de kräver vad gäller särskild behörighet.

lysen inte är begränsad till de elever som sökte till HHS utan kan ses som ett mer generellt exempel på effekten av en betygsreform för elever som kan söka attraktiva utbildningar och som bara en mindre andel av en elevkohort kan komma in på.

Data och metod

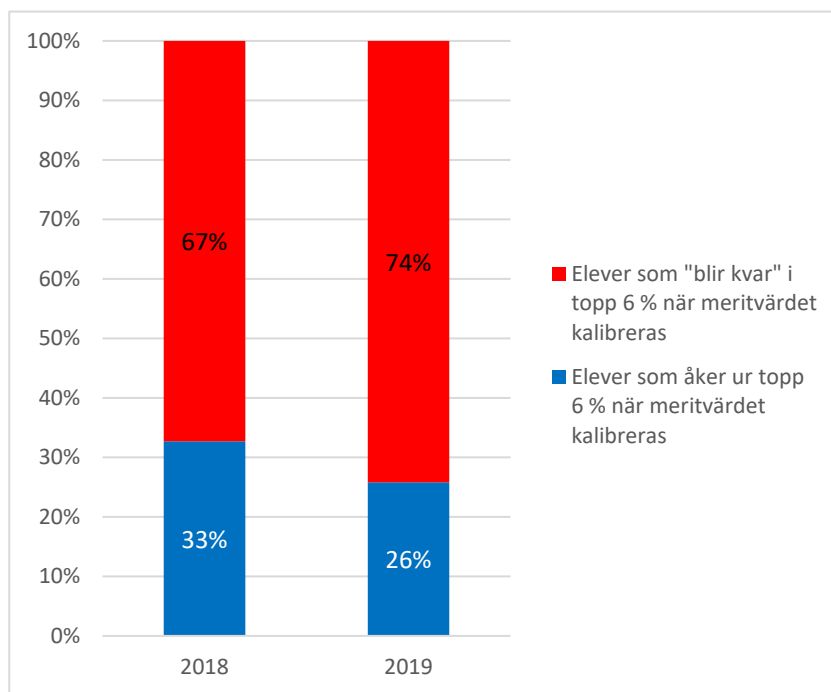
Vi utgår i denna analys från samma dataunderlag som tidigare i avsnitt 1.5. I det första steget av analysen identifieras hur stor andel av eleverna som uppfyller kravet på ett meritvärde på minst 19,17 för respektive år. För 2018 är andelen 5,64 % och för 2019 är den 6,29 %. Utifrån dessa andelar identifierar vi samma andelar för det kalibrerade meritvärdet som mäts på en annan skala (årsvis standardiserad med medelvärde 50 och standardavvikelse 15). Slutligen identifierar vi med hjälp av dessa andelar det antagningsvärde på skalan för det kalibrerade meritvärdet som motsvarar 19,17 på originalskalan för det okalibrerade meritvärdet. Vi kan då identifiera de elever som uppfyller respektive inte uppfyller antagningsgränsen för det ursprungliga meritvärdet och det kalibrerade meritvärdet.

Resultat

Analysen i figur 1.39 omfattar endast de elever som hade ett okalibrerat meritvärde (jämförelsetal) på minst 19,17 poäng. Det innebär 841 elever 2018 respektive 910 elever 2019 (i dataunderlaget för analysen). Detta antagningskrav innebär att eleverna måste tillhöra de 6 % högst presterande eleverna enligt det okalibrerade meritvärdet. Dessa elever motsvarar således 100 % för respektive år i figuren. Av dessa elever uppfyllde 67 %, även kravet för antagning enligt det kalibrerade meritvärdet.⁸ Samtidigt uppfyllde 33 %, *inte* kravet enligt det kalibrerade meritvärdet. För år 2019 är motsvarande andel som inte uppfyller kravet 26 %. Detta får anses som ganska stora andelar och indikerar att en betygsreform med kalibrerade meritvärden kommer att få reell betydelse för vilka som antas till de mest attraktiva utbildningarna.

⁸ 19,17 poäng enligt det okalibrerade meritvärdet motsvarar de 6 % högst presterande eleverna. Att uppfylla intagningskravet med det kalibrerade meritvärdet motsvaras därför av att tillhöra de 6 % högst presterande eleverna enligt det kalibrerade meritvärdet.

Figur 1.39 Förändring i elevsammansättning före och efter meritvärdeskalibrering för de 6 % högst presterande eleverna



Källor: Utredningens egna beräkningar utifrån betygsstatistik från Skolverket och högskoleprovsstatistik från Umeå universitet.

På vilka sätt skiljer sig de elever som kom in på HHS enligt det okalibrerade meritvärdet men inte enligt det kalibrerade meritvärdet från de elever som klarade antagningen enligt båda kriterierna? Och vilka är skillnaderna jämfört med de elever som inte antas enligt meritvärdet (jämförelsetalet) men som skulle antas enligt det kalibrerade meritvärdet?

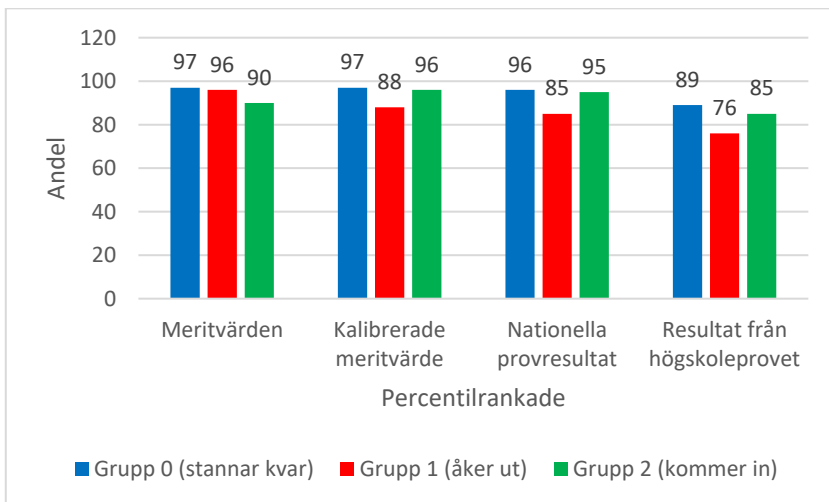
I figur 1.40 har vi delat upp eleverna i fyra kategorier:

- 0 – blir kvar: De elever som uppfyller kraven enligt såväl det okalibrerade som det kalibrerade meritvärdet.
- 1 – åker ut vid meritvärdeskalibrering: De elever som uppfyller kravet enligt meritvärdet men inte enligt det kalibrerade meritvärdet.

- 2 – kommer in med meritvärdeskalibrering: De elever som inte uppfyller kravet enligt meritvärdet men däremot enligt det kalibrerade meritvärdet.
- 3 – aldrig aktuella: De elever som inte uppfyller kraven enligt vare sig meritvärdet eller det kalibrerade meritvärdet.

Vi fokuserar på grupp 1 i förhållande till grupp 0 och 2, dvs. de som ”åker ur” utbildningen utifrån det kalibrerade meritvärdet i jämförelse med de som ”blir kvar” respektive de som ”kommer in”. För att kunna jämföra olika mått är det fördelaktigt att titta på de percentilrankade måtten eftersom dessa är jämförbara. I figur 1.40 redovisas de mest centrala resultaten i form av stapeldiagram.

Figur 1.40 Percentilrankade meritvärden, kalibrerade meritvärden, nationella provresultat respektive resultat från högskoleprovet för elevgrupp 0, 1 och 2



Anm.: Grupp 0 (blir kvar): De elever som uppfyller kraven enligt såväl det okalibrerade som det kalibrerade meritvärdet.

Grupp 1 (åker ut vid meritvärdeskalibrering): De elever som uppfyller kravet enligt meritvärdet men inte enligt det kalibrerade meritvärdet.

Grupp 2 (kommer in med meritvärdeskalibrering): De elever som inte uppfyller kravet enligt meritvärdet men däremot enligt det kalibrerade meritvärdet.

Källor: Utredningens egna beräkningar utifrån betygsstatistik från Skolverket och högskoleprovsstatistik från Umeå universitet.

Jämfört med grupp 0 är det percentilrankade (okalibrerade) meritvärdet för grupp 1 endast en percentilenhet lägre och sex percentilenheter högre än grupp 2. Med det kalibrerade meritvärdet är cirka nio percentilenheter lägre än grupp 0 och åtta percentilenheter lägre än grupp 2. Detta är förväntat eftersom det är just det kalibrerade meritvärdet som utgör den nya antagningsgrunden.

Vi kan också se att elevgrupp 1 (de som åker ut) har lyckats betydligt sämre på de nationella proven jämfört med både de som blir kvar och de som kommer in (grupp 0 och 2). Skillnaden uppgår till 11 respektive 10 percentilenheter. Kan det vara så att denna grupp representerar elever som har "haft en dålig dag på provet" och som läraren enligt gällande betygssättningsregelverk kanske bedömde kunde betydligt mer än de lyckades visa på provet? Så skulle det rent teoretiskt kunna vara.

Eftersom alla elever i underlaget också har skrivit högskoleprovet, vilket i dessa sammanhang kan ses som ett oberoende mått på elevers kunskaper, kan vi se om elevgrupp 1 bara hade otur på de nationella proven. Enligt figuren skiljer sig resultaten på högskoleprovet dock ännu mer mellan grupperna än på de nationella proven, 13 percentilenheter lägre resultat jämfört med grupp 0 respektive 9 elevpercentiler sämre jämfört med grupp 2.

Det tolkar vi som att dessa elever (grupp 1) knappast bara har haft "otur" på proven utan i stället är en grupp elever som fått betyg som inte motsvarar deras kunskaper (relativt övriga elever). När meritvärdena kalibreras så att de mäts på samma skala elimineras olikvärdig betygssättning mellan skolor och vi erhåller en mer rättvisande bild av elevernas kunskaper. Konsekvensen blir att en elevgrupp åker ut och en annan elevgrupp kommer in.

Detta ska inte tolkas som att den grupp som kommer in fått mer rättvisande betyg från början. Det går inte att säga om en elevgrupp fått en restriktiv eller generös betygssättning i *absolut* mening. Vad vi däremot kan säga är om en viss grupp fått *relativt* generösa eller *relativt* restriktiva betyg. Det viktiga att förstå är att de har betygssatts på olika skalor, sannolikt till stor del för att lärare på olika skolor haft olika elevunderlag och ställt olika krav för ett visst betyg. En och samma prestation har helt enkelt inte genererat samma betyg på olika skolor. När vi kontrollerar för denna olikvärdighet och mäter elevernas kunskaper på en gemensam skala får vi ett mer rättvisande utfall.

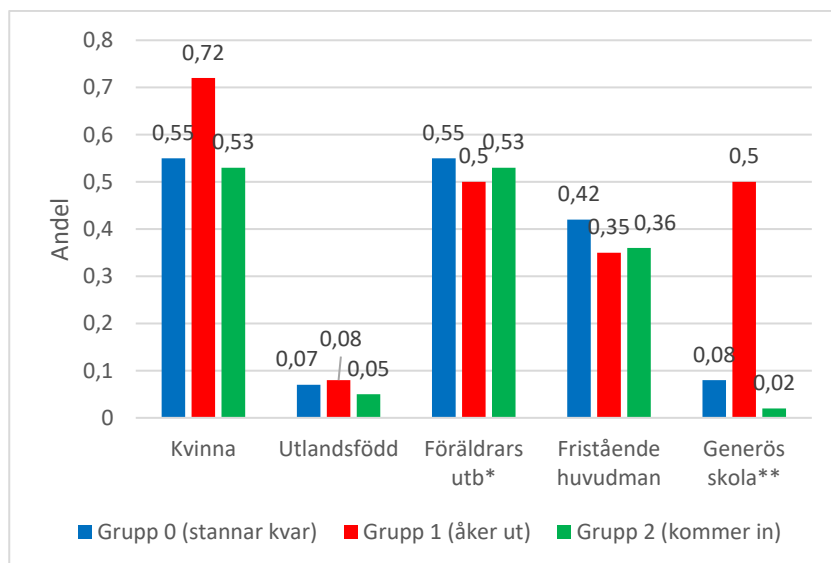
Om gruppen som åker ut har fått för generösa betyg eller om det är gruppen som kommer in som har fått för restriktiva betyg kan vi alltså inte svara på och frågan saknar egentligen relevans. Det viktiga är att vi får ett mer rättvisande mått på elevernas kunskaper och att det är detta som antagningen baseras på. Resultaten på de nationella proven tillsammans med resultatet på högskoleprovet utgör starka indicier på att det kalibrerade meritvärdet genererar en mer rättvis och rättvisande rangordning av elevernas kunskaper jämfört med det ursprungliga okalibrerade meritvärdet.

Elevsammansättning med avseende på demografi och skolkategori

I figur 1.41 jämförs elevgrupp 0, 1 och 2 med avseende på elevsammansättning utifrån andel kvinnor, utlandsfödda, föräldrars genomsnittliga utbildningsnivå, typ av huvudman samt om skolan definieras som ”generös” i betygssättningen (relativt andra skolor). För att en skola ska kategoriseras som ”generös skola” ska den ha tappat minst 10 percentilenheter i skolrankningen med kalibreringen. Samtliga värden i figur 1.41 kan tolkas som andelar förutom föräldrars genomsnittliga utbildningsnivå som är ett genomsnitt på en kontinuerlig skala.

Det framgår att andelen kvinnor är betydligt större i den elevgrupp som ”åker ur” (grupp 1). Vad gäller andelen utlandsfödda är den något större (1 procentenhet) i elevgrupp 1 jämfört med elevgrupp 0. Däremot är andelen utlandsfödda klart lägre i gruppen som ”kommer in” (grupp 2). När det gäller föräldrars utbildningsnivå är den lägre i gruppen som åker ur. Andelen elever i fristående skolor är högre i gruppen som ”stannar kvar” (grupp 0). Slutligen är andelen elever som kommer från ”generösa” skolor betydligt högre i elevgrupp 1, de som ”åker ur” jämfört med övriga grupper.

Figur 1.41 Elevsammansättning med avseende på demografiska variabler



Anm.: Kvinna = 1 för kvinna. Utlandsfödd = 1 för utlandsfödd. Friskola huvudman = 1 för om elev gått på en fristående skola.

*Variabeln [edu13/10] = föräldrarnas genomsnittliga utbildningsnivå på en 13-gradig skala dividerat med 10 för att passa in i stapeldiagrammet.

**Generös skola = 1 om skolan förlorat minst 10 percentilenheter på skolrankningen i samband med kalibreringen.

Källor: Utredningens egna beräkningar utifrån Skolverkets statistik.

Sammantaget framgår alltså att kvinnor, utlandsfödda, elever vars föräldrar har en relativt lägre utbildningsnivå och elever som kommer från skolor som sätter relativt generösa betyg är överrepresenterade i gruppen som åker ut jämfört med i övriga grupper. Detta mönster syns delvis även i analysen av hela elevkohorten (inte bara de som kommer in på HHS).

Att kvinnor är överrepresenterade i elevgruppen som "åker ur" ska inte nödvändigtvis ses som att de skulle vara gynnade i betygs-sättningen i dagens system. En annan orsak är att de är överrepresenterade på gymnasieprogram där det är lättare att få höga betyg, såsom SA, EK och ES.

Vad gäller fristående huvudman ska läsaren återigen uppmärksammas på analysens begränsning att dessa analyser bygger på provresultat som lärarna/skolorna själva har rättat. Det finns forskning

som visar att proven i genomsnitt rättas mer generöst i fristående skolor jämfört med i kommunala skolor.⁹

1.3.9 Konsekvenser av "Strannegårdskriteriet"

Handelshögskolan i Stockholm (HHS) tillämpar fr.o.m. 2025 ett tilläggskrav för behörighet till utbildningen på HHS som innebär att eleverna förutom ett visst meritvärde också måste ha ett resultat på minst 1,25 poäng på högskoleprovet.

Hur stor andel av de sökande kommer att klara av det tilläggskravet? I föregående avsnitt såg vi att den genomsnittliga nivån på resultaten från högskoleprovet var betydligt lägre i den grupp som åkte ut när vi använde det kalibrerade meritvärdet i stället för originalmeritvärdet. Motsvarande undersöks i detta avsnitt för högskoleprovet.

Data och metod

Metoden är i stort sett densamma i denna analys som tidigare, men den skiljer sig på ett par punkter. Vi urskiljer gruppen som uppfyller antagningskravet enligt meritvärdet och sen undersöker vi hur stor andel av dessa elever som inte uppfyller kravet på 1,25 på högskoleprovet. En skillnad är att vi inte specifikt har undersökt de elever som ersätter dem som åker ur i denna analys (vilket vi skulle kunna ha gjort men då sänks antagningsgränsen för meritvärdet).

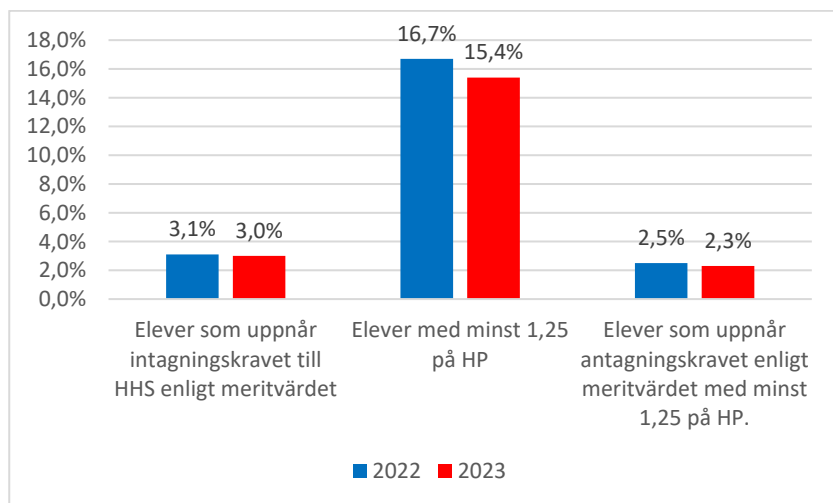
Vi har använt data från 2022 och 2023, vilka är de senast tillgängliga. Det innebär att eleverna har gått ut gymnasieskolan 2022 eller 2023 från de fem vanligaste högskoleförberedande programmen (EK, ES, NA, SA, TE). För att ingå i analysen ska eleverna också ha genomfört högskoleprovet samma år som de gick ut gymnasieskolan eller året innan. Dessa restriktioner innebär att analyspopulationen uppgår till 23 103 respektive 29 570 elever för 2022 respektive 2023. Uppgifter om nationella provresultat är tyvärr begränsade för dessa elevkohorter, speciellt vad gäller matematik 1a/1b/1c som många av dessa elever skulle ha gjort under pandemin och som vi behöver för att beräkna kalibrerade meritvärden. Därmed blir analysen något mer begränsad än den i föregående avsnitt.

⁹ Se Tyrefors Hinnerich & Vlachos (2012, 2013).

Resultat

Antagningspoängen vad gäller meritvärde till HHS 2022 respektive 2023 var 19,43 respektive 19,58. I figur 1.42 visas hur stor andel av eleverna som klarade antagningskravet enligt meritvärdet, hur stor andel som hade minst 1,25 på högskoleprovet samt hur stor andel som klarade båda kraven.

Figur 1.42 Andel elever som klarar intagningskrav till HHS enligt meritvärde, 1,25 på högskoleprovet samt kombinationen av båda kraven, 2022 och 2023

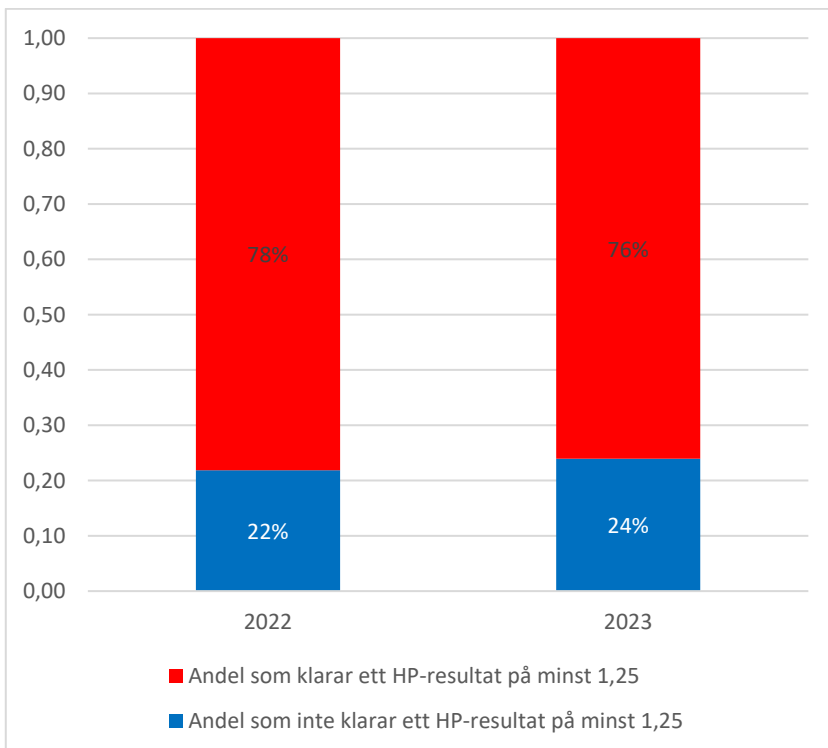


Källor: Utredningens egna beräkningar utifrån Skolverkets och Umeå universitets statistik.

Ungefär 3 % av eleverna hade ett meritvärde på minst den faktiska antagningspoängen till HHS 2022 och 2023. Andelen som hade minst 1,25 på högskoleprovet uppgick till 15–17 %. Andelen elever som uppfyllde båda kraven uppgick till ungefär 2,5 %.

I figur 1.43 redovisas hur stor andel av eleverna som hade tillräckligt högt meritvärde men som inte hade ett högskoleprovsresultat på minst 1,25. År 2022 var det 22 % av dem som uppfyllde meritvärdeskravet som inte nådde upp till 1,25 poäng på högskoleprovet. År 2023 var motsvarande andel 24 %. Det innebär att även en så pass låg gräns på högskoleprovet medför betydande förändringar i elevsammansättningen.

Figur 1.43 Andel elever som uppfyller respektive inte uppfyller ett högskoleprovsresultat på minst 1,25 av de elever som har ett meritvärde på minst 19,58 respektive 19,43 för år 2022 respektive 2023



Källor: Utredningens egna beräkningar utifrån Skolverkets och Umeå universitets statistik.

1.4 Meritvärdeskalibrering vid specialfall

I detta avsnitt behandlas ett antal specialfall för vilka den generella meritvärdeskalibreringen kan behöva anpassas för att få ett optimalt utfall. Ett sådant exempel är skolenheter med mycket få elever. Ett annat exempel är hur meritvärden för elever som inte deltar på alla prov ska beräknas. Slutligen analyseras hur elever som sticker ut med extrema resultat, s.k. *outliers*, ska hanteras. Det sistnämnda gäller framför allt elever som riskerar att drabbas otillbörligt på grund av takeffekter i betygssättningen.

1.4.1 Små skolor

För att få en uppfattning om hur mycket mätsäkerheten minskar när antalet elever på skolan minskar kan man analysera korrelationen på skolnivå mellan en skolas genomsnittliga ”sanna kunskap” och motsvarande genomsnittliga betyg/resultat utifrån olika betygsmodeller.¹⁰ Detta måste göras utifrån en datasimulering och är här baserad på en ämnesvis simulering. Det innebär att det är skolans genomsnittliga ”sanna kunskap” i ett enskilt ämne som korreleras med skolans genomsnittliga betyg/provresultat i samma ämne.

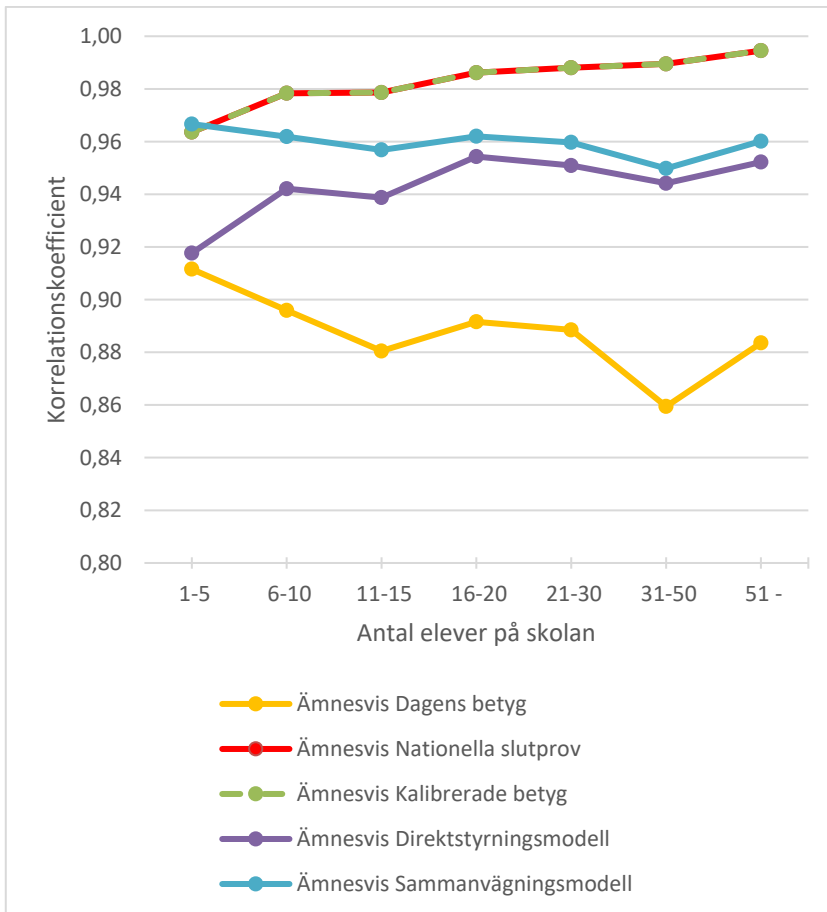
I figur 1.44 visas korrelationskoefficienterna för olika betygsmodeller när de korreleras mot skolans genomsnittliga ”sanna kunskap” i ett givet ämne. Meritvärdeskalibreringsmodellen kalibrerar inte enskilda ämnen men kan i denna analys likställas med att kalibrera ämnesvis.

Generellt är korrelationen högst för modellen som kalibrerar betygen ämnesvis samt modellen med nationella slutprov. I figuren syns endast linjen för den ämnesvisa kalibreringsmodellen därför att de nationella slutproven har samma värden. Detta är ingen slump utan ett resultat av att den ämnesvisa kalibreringsmodellen innebär att skolans genomsnittliga betyg blir lika med skolans genomsnittliga provresultat från det nationella slutprovet. Detta innebär att om man vill mäta en skolas ”sanna kunskap” med ett så litet mätfel som möjligt, är de nationella provresultaten det bästa vi har att tillgå, särskilt om de rättas centralt.

Vi kan också se att mätsäkerheten är något lägre när elevantalet är litet jämfört med när elevantalet är stort. För skolor med mer än 50 elever är korrelationen med skolans sanna genomsnittliga kunskap 0,995 medan den för skolor med högst fem elever är 0,964. Observera dock att även för dessa extremt små skolor så är korrelationen högre än för dagens betygssystem som endast kommer upp till 0,912. Även direktstyrningsmodellen har betydligt lägre korrelation med skolans ”sanna kunskap” när elevantalet är så pass litet. Vi kan också se att för extremt små skolor uppvisar sammanvägningsmodellen lika bra korrelation som proven.

¹⁰ Genom att analysen baseras på simulerade eller hypotetiska data så kan vi tillskriva varje elev ett värde som per definition är elevens ”sanna kunskap” och som sedan respektive betygssystem kan jämföras med. I ett perfekt betygssystem får eleven ett betyg som exakt motsvarar elevens ”sanna kunskap”. Eller annorlunda uttryckt, korrelationen mellan en rangordning av elevens ”sanna kunskap” och en rangordning av elevens betyg är 1,0 i ett perfekt betygssystem.

Figur 1.44 Korrelation mellan skolors genomsnittliga "sanna kunskap" och olika resultatmått, uppdelat efter skolans storlek



Anm.: Nationella slutprov och kalibrerade betyg ämnesvis har båda samma värden och syns därför som en randig linje i figuren. Analysen bygger på simulering på ämnesnivå (matematik), genomsnitt 2013–2023.
 Källa: Utredningens egna beräkningar utifrån Skolverkets statistik.

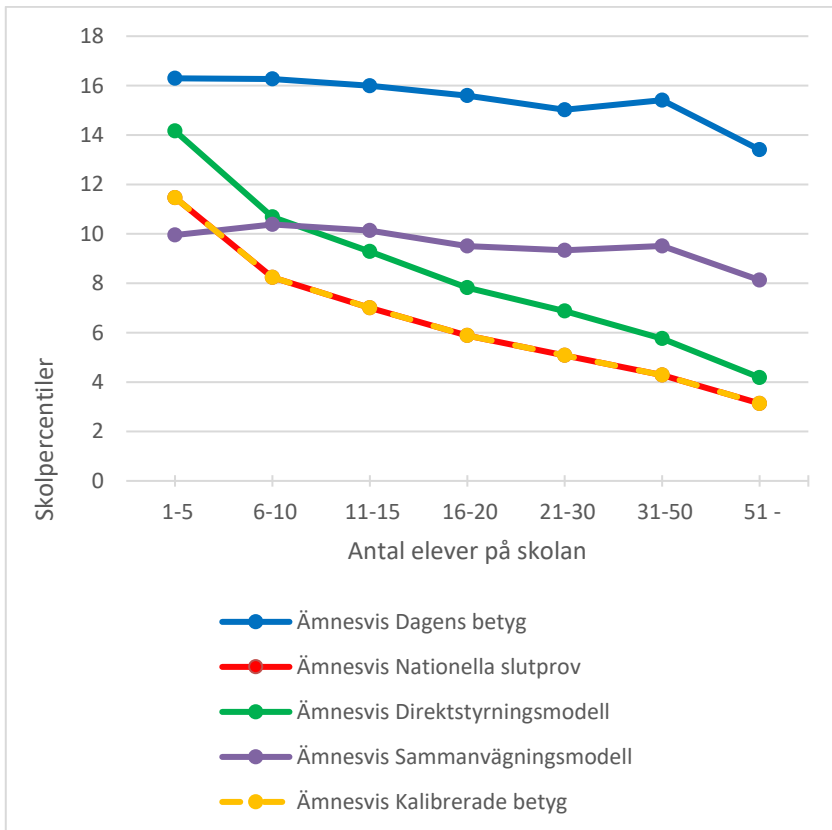
I figur 1.45 visas den genomsnittliga avvikelser i percentilrankning mellan en skolans genomsnittliga "sanna kunskapsnivå" och motsvarande genomsnitt enligt olika betygsmodeller. Ett högre värde innebär en större skillnad i skolans rankning enligt betygsmodellen och den rankning som skolan borde ha enligt skolans "sanna kunskap". Ett lägre värde innebär således en bättre överensstämmelse med den "sanna kunskapsnivån", dvs. en mer likvärdig modell.

Dagens betygssystem uppvisar sämst passning medan nationella slutprov, och därmed även modellen med ämnesvis kalibrering, sett över hela figuren klarar sig bäst. För skolor med högst fem elever uppvisar sammanvägningsmodellen dock något bättre samstämmighet med den ”sanna kunskapsnivån”.

Båda figurerna, 1.44 och 1.45, ger samma bild, det är bara två olika sätt att analysera sambandet mellan skolans sanna genomsnittliga kunskapsnivå och olika betygsutfall. Analyserna visar att det inte finns någon anledning att låta skolenheter med högst fem elever beräkna meritvärden helt och hållet baserat på lärarsatta betyg, dvs. dagens betygssystem eftersom det ger ett sämre likvärdighetsutfall enligt simuleringen.¹¹ Däremot visar analyserna att en sammanvägningsmodell har vissa förtjänster för sådana skolor eftersom den uppvisar ett något, om än marginellt, bättre utfall jämfört med en ämnesvis kalibreringsmodell för just kategorin skolor med upp till fem elever.

¹¹ Observera att simuleringen använder ingångsvärden som baseras på empiriska data. Det innebär t.ex. att när vi simulerar betygen enligt dagens system så inkluderas effekter av olikvärdig betygssättning, vilka baseras på vad analyser på riktiga data har påvisat. Se Skolverket (2020) för en mer ingående beskrivning av hur en realistisk simuleringsmodell är uppbyggd.

Figur 1.45 Genomsnittlig avvikelse i skolors percentilrankning mellan skolors genomsnittliga "sanna kunskap" och olika resultatmått – uppdelat efter skolans storlek



Anm.: Nationella slutprov och ämnesvis kalibrerade betyg har båda samma värden och syns därför som en rödgulrandig linje i figuren. Analysen bygger på simulering på ämnesnivå (matematik), genomsnitt 2013–2023.

Källa: Utredningens egna beräkningar utifrån Skolverkets statistik.

1.4.2 Elever som inte deltar på samtliga prov

Elever som inte deltar på samtliga nationella slutprov kan ändå erhålla ett kalibrerat betygsvärde. Dessa elevers betygsvärde kalibreras efter samma kalibreringsalgoritm som de elever som deltagit på alla slutprov.

Det är dock viktigt att förstå att själva kalibreringsalgoritmen enbart bestäms utifrån det genomsnittliga provresultatet respektive det genomsnittliga betygsvärdet för de elever som deltagit på samt-

liga slutprov. En skolas elever kan därmed inte vinna eller förlora på om en relativt låg- eller högpresterande elev inte deltar på provet. Om t.ex. en lågpresterande elev inte deltar kommer skolans genomsnittliga provresultat visserligen att bli högre men det gäller även skolans genomsnittliga betygsvärde. Eftersom det avgörande för kalibreringen är relationen mellan skolans genomsnittliga provresultat och genomsnittliga betygsvärde finns inga fördelar för övriga elever om denna elev inte deltar på provet. Däremot kommer eleven själv att förlora på att inte delta på provet eftersom den får ett provresultat på 0 poäng på det prov som eleven inte deltagit på. Detta ligger sedan till grund för den slutliga meritvärdesberäkningen där elevens individuella provresultat väger 30 %. Det finns därför starka incitament för samtliga elever, oavsett prestationsnivå, att delta på slutproven. I tabell 1.8 och figur 1.46 visas principen för hur kalibreringen fungerar när en elev inte har deltagit på ett slutprov.

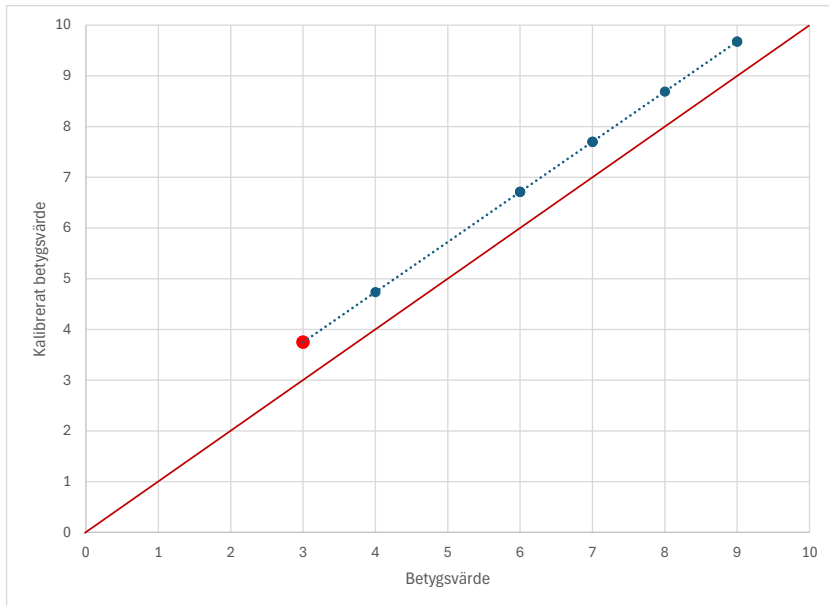
Tabell 1.8 Sifferunderlag för exempel med kalibrering när vissa elever inte deltagit på samtliga nationella slutprov

Elev	Betygs-värde	Prov-resultat	Kalibrerat betygsvärde	Meritvärde
1	9	10	9,7	9,8
2	9	9	9,7	9,5
3	8	9	8,7	8,8
4	7	8	7,7	7,8
5	7	8	7,7	7,8
6	7	7	7,7	7,5
7	6	7	6,7	6,8
8	6	7	6,7	6,8
9	6	6	6,7	6,5
10	4	5	4,7	4,8
11*	3	0	3,7	2,6
Medelvärde	6,90	7,60	7,60	7,60
Spridning (sd)	1,4	1,4	1,4	1,4

* Observera Bortfallselevens (nr 11) betygsvärde respektive kalibrerade betygsvärde ingår vare sig i gruppens medelvärde eller i beräkningen av standardavvikelsen (sd).

Källa: Utredningens egna beräkningar utifrån Skolverkets statistik.

Figur 1.46 Kalibrering av betygsvärde för elever som deltagit respektive ej deltagit på ett slutprov



Anm.: Bortfallselevens betygsvärde påverkar ej övriga elevers kalibrering. Däremot avgörs bortfallselevens kalibrerade betygsvärde av skolans genomsnittliga provresultat respektive genomsnittliga betygsvärde utan hänsyn till bortfallselevens betygsvärde.

Källa: Utredningens egna beräkningar.

I detta hypotetiska exempel består skolan av 11 elever, varav 10 har deltagit på slutprovet. I exemplet har vi för enkelhetens skull antagit att det bara finns ett enda slutprov. Samtliga genomsnitt för skolan baseras endast på de 10 elever som deltagit på provet. Som exemplet är konstruerat är skolans genomsnittliga provresultat 7,6 poäng för de 10 deltagande elevernas medan det genomsnittliga betygsvärdet för samma 10 elever är 6,9 poäng. Detta innebär att skolans genomsnittliga betygsvärde ska kalibreras upp från 6,9 till 7,6 poäng. För enkelhetens skull har vi i detta exempel även antagit att spridningen (standardavvikelsen) är identisk för provresultat och betygsvärde, vilket innebär att spridningen inte behöver kalibreras. Därmed ska varje enskild elevs kalibrerade betygsvärde kalibreras upp så att det blir 0,7 poäng högre än deras betygsvärde.

Bortfallselevens betygsvärde är tre poäng, men eftersom skolan i genomsnitt ska höjas med 0,7 poäng kommer även bortfallselevens

kalibrerade betygsvärde att höjas med 0,7 poäng. Vi tillämpar helt enkelt samma kalibreringsalgoritm på bortfallseleven som på de elever som deltagit på provet.

Det viktiga att observera är att denna kalibreringsalgoritm enbart är beräknad utifrån de elever som deltagit på provet. Det finns ingen anledning att tro att eleverna som deltog på provet skulle ha tjänat på att bortfallseleven deltagit. Bortfallseleven kan förväntas ha påverkat skolans genomsnittliga provresultat lika mycket som skolans genomsnittliga betygsvärde och därmed hade kalibreringsalgoritmen blivit den samma. Om eleven är relativt lågpresterande som i detta exempel, kommer eleven visserligen att förväntas sänka skolans genomsnittliga provresultat genom att delta, men eleven kommer också att sänka skolans genomsnittliga betygsvärde genom att delta. I verkligheten kan det naturligtvis alltid finnas en slumpvariation som gör att eleven hade kunnat prestera lite bättre eller lite sämre än förväntat på provet, men det finns ingen systematisk bias i detta tillvägagångssätt. Det innebär att den prickade linjen som beskriver kalibreringen i relation till den streckade referenslinjen (som beskriver en situation där det kalibrerade betygsvärdet är identiskt med det ursprungliga betygsvärdet) kan förväntas vara densamma även om eleven hade deltagit på provet.

I det sista steget ska meritvärdet beräknas som ett viktat genomsnitt av elevens kalibrerade betygsvärde och elevens individuella provresultat. I detta steg sätts bortfallselevens provresultat till 0 poäng och meritvärdet beräknas följaktligen som: $0,70 \cdot 3,7 + 0,30 \cdot 0,0 = 2,6$ poäng (se tabell 1.8 och figur 1.46).

När eleven deltagit på några men inte samtliga slutprov

I exemplet ovan antas att det bara finns ett slutprov. Om vi i stället har en situation med fem slutprov och eleven har deltagit på fyra av dessa kommer fortfarande kalibreringsalgoritmen att baseras på *endast* de elever som deltagit på samtliga fem slutprov. Däremot kommer bortfallselevens samlade provresultat inte att vara 0 utan ett genomsnitt av de fyra provresultaten och 0 för det prov som eleven inte deltagit på. Anta t.ex. att bortfallseleven har provresultaten 2, 3, 4 och 6 poäng för de fyra prov som eleven deltagit på. Då beräknas ett genomsnittligt provresultat som $(2 + 2 + 4 + 6 + 0)/5 = 2,8$ poäng.

Om eleven har samma kalibrerade betygsvärde som i exemplet ovan blir nu meritvärdet i stället: $0,70 \cdot 3,7 + 0,30 \cdot 2,8 = 3,43$ poäng.

När en stor andel elever inte genomför proven

Om andelen elever som inte deltar på samtliga prov blir stort finns risk för att relationen (som bestämmer kalibreringsalgoritmen) mellan skolans genomsnittliga provresultat och betygsvärde inte är representativ för den elevgrupp som inte deltagit på alla prov. Om dessa bortfallselever kan betraktas som slumpmässigt dragna ur hela elevpopulationen på skolan är det inget problem men om bortfallet är systematiskt kan kalibreringen för dessa bortfallselever generera felaktiga kalibrerade betygsvärden.

En sådan situation kan naturligtvis uppstå redan vid mindre bortfallsandelar men konsekvenserna på systemnivå kommer att vara små. Det finns därför anledning att sätta en maxgräns för andelen bortfallselever på en skola för att kalibreringsprocessen ska kunna tillämpas med bibehållen kvalitet. En sådan maxgräns är till viss del godtycklig men vår bedömning är att den bör sättas vid 50 %. Detta innebär att om andelen elever som inte deltagit på samtliga slutprov överstiger 50 % (på skolenheten) kan dessa bortfallselevens betygsvärde ej kalibreras på ett tillförlitligt sätt och därmed kan inget slutligt meritvärde beräknas för dessa elever.

Observera att denna maxgräns inte på något sätt påverkar de elever som har deltagit på samtliga prov. För provdeltagande elever kalibreras betygsvärdena som om skolan endast bestod av dessa elever (något som också är sant när andelen bortfall är mindre än 50 %). Maxgränsen är viktig inte bara för den statistiska kvaliteten i kalibreringen utan även som en signal till skolan att verka för att bortfallet minimeras.

1.5 Simulering av betygsmodeller – uppdaterade resultat

I detta avsnitt redovisas några övergripande resultat från en simuleringsstudie av olika betygsmodeller. Simuleringsstudien bygger i huvudsak på den studie som Skolverket redovisade i *Likvärdiga betyg*

och meritvärden.¹² Skillnaden mot Skolverkets simuleringsstudie är att den här har uppdaterats med senare data (2013–2023) samt att den har kompletterats med en simulering av en mer fingradig betygsskala för att se vilken effekt det får. I simuleringen valdes en skala 1–20.¹³

1.5.1 Grundläggande principer för simulering

Att simulera data innebär i princip att data är hypotetiska eller påhittade om man så vill. Fördelen med detta är att vi i analysen har full kontroll, vilket i detta fall specifikt innebär att vi per definition känner till elevernas ”sanna” kunskap. I simuleringen ges sedan eleverna ett provresultat och ett betyg, baserat på deras ”sanna kunskap” men också med ett visst mätfel samt även andra systematiska effekter som bl.a. beror på vilken skola och vilken lärare eleven har. Simuleringen utförs med olika betygsmodeller som ska efterlikna den situation som karaktäriserar den specifika betygsmodellen. Om simuleringsmodellen görs tillräckligt realistisk kan utfallet från olika betygsmodeller jämföras på ett meningsfullt sätt och vi kan dra slutsatser om vilka betygsmodeller som fungerar bättre respektive sämre ur ett likvärdighetsperspektiv. Detta görs genom att jämföra elevernas rangordning (percentilrankning) enligt en viss betygsmodell med motsvarande rangordning enligt elevernas ”sanna kunskap”. Om betygsmodellen är perfekt likvärdig kommer skillnaden i elevernas rangordning enligt betygsmodellen och rangordningen enligt deras ”sanna kunskap” att vara noll. Ju större skillnaderna är i rangordning när vi tar hänsyn till alla elever i årskullen, desto mindre likvärdig är modellen.

Simuleringen är baserad på grundskolans 10 ämnen med nationella prov men genomförs ämne för ämne.¹⁴ Detta innebär att vi inte kan redovisa resultat från en simulering där hela meritvärdet kalibreras enligt den meritvärdeskalibreringsmodell som utredningen föreslår. Icke desto mindre kan resultaten från denna ämnesvisa simulering ändå ge mycket goda indikationer på hur olika betygsmodeller

¹² För en närmare redogörelse för själva simuleringsmodellens konstruktion, se Skolverket (2020).

¹³ När denna simuleringsstudie genomfördes var en 1–20-skala ett av flera tänkbara alternativ som utreddes.

¹⁴ Att genomföra en simuleringsstudie för samtliga ämnen på en gång är betydligt mer komplicerat och utredningen har av tidsskäl avgränsat analysen enligt vad som redovisas här. Simuleringens resultat går hursomhelst att överföra även till fallet med meritvärdeskalibrering då betygsmodellernas egenskaper kommer att gälla även för meritvärdet.

fungerar respektive inte fungerar. Detta gäller naturligtvis under förutsättning att simuleringen är uppbyggd på ett realistiskt sätt, vilket innebär att de ingångsparametrar som används i simuleringen är kalibrerade på ett sådant sätt att de motsvarar förhållanden som i största möjliga utsträckning botten i empiriska observationer.

Även om en simuleringsstudie i princip är baserad på hypotetiska ingångsdata ska data vara så pass realistiska att det som kommer ut ur simuleringen, själva utfallet, säger oss någonting om hur likvärdigheten skulle påverkas om vi tillämpade en viss betygsmodell jämfört med en annan.

1.5.2 Betygsmodeller

De betygsmodeller som utvärderats i denna simuleringsstudie är:

1. Dagens betygssystem (Nollmodellen).
2. Examensprovsmodellen.
3. Direktstyrningsmodellen.
4. Kombinationsmodellen (oberoende slutprov).
5. Sammanvägningsmodellen.
6. Modereringsmodellen.

Dagens betygssystem innebär att läraren sätter betyg utan några restriktioner. Läraren förväntas ta viss hänsyn till resultatet på det nationella provet men denna hänsyn varierar mellan olika lärare och är inte på något sätt juridiskt bindande.

Examensprovsmodellen innebär att läraren är helt frikopplad från betygssättningen och elevens betyg i ämnet helt och hållet avgörs av elevens resultat på det nationella slutprovet.

Direktstyrningsmodellen innebär att lärarna på varje skola måste anpassa sina betyg i ett givet ämne så att skolans genomsnitt motsvarar skolans genomsnitt på det nationella slutprovet. På elevnivå är det upp till varje lärare att avgöra hur mycket den enskilda elevens provresultat ska påverka elevens individuella betyg.

Kombinationsmodellen med oberoende slutprov innebär samma som ovan men med den skillnaden att läraren sätter betyg på eleven utan att känna till den enskilda elevens provresultat. Lärarna på skolan

måste däremot fortfarande sätta betyg så att skolans genomsnittliga betygsnivå motsvarar skolans genomsnittliga provresultat (som ovan). Elevens individuella provresultat läggs sedan till elevens lärarsatta betyg för att bilda det slutliga betyget i ämnet.

Sammanvägningsmodellen innebär att lärarna sätter betyg oberoende av elevernas resultat på det nationella slutprovet, på både elev- och skolnivå. Det senare innebär att lärarna inte behöver anpassa sina betyg efter hur skolan i genomsnitt presterar på det nationella slutprovet utan kan sätta betyg på samma sätt som i dagens betygssystem men utan information om hur eleverna presterat på slutprovet. Ett genomsnitt skapas av lärarbetyget och elevens provresultat och utgör det slutliga betyget i ämnet.

Modereringsmodellen innebär att lärarna sätter betyg helt oberoende av elevernas resultat på det nationella slutprovet. Dessa betyg justeras sedan ämnesvis på central nivå så att skolans genomsnittliga betyg i respektive ämne motsvarar skolans genomsnittliga provresultat. Därefter beräknas ett viktat genomsnitt av elevens modererade betyg och elevens individuella provresultat som utgör det slutliga betyget i ämnet.

1.5.3 Resultat

I figur 1.47 redovisas resultaten från simuleringen på elevnivå. Varje betygsmodell representeras av en stapel som visar hur olikvärdig betygsmodellen är. En högre stapel indikerar således en sämre betygsmodell ur ett likvärdighetsperspektiv. En perfekt likvärdig betygsmodell skulle ha värdet noll; detta uppnås inte med någon modell. Det beror på att det mått som vi använder inte bara fångar upp systematisk olikvärdighet, t.ex. att olika skolor sätter mer eller mindre generösa betyg. Måttet fångar dessutom upp mätfel som uppstår på grund av att ett prov, en inlämningsuppgift eller en muntlig presentation aldrig helt perfekt kan mäta en elevs ”sanna kunskap” i ämnet. Därmed har inte heller läraren perfekt information om elevens ”sanna kunskap”.¹⁵

¹⁵ För en mer grundlig genomgång av såväl simuleringsmodellen samt måttet som används hänvisas till Skolverket (2020).

Figur 1.47 är uppdelad i två sektioner. I den vänstra sektionen redovisas de sex betygsmodellerna baserat på en 6-gradig F–A skala.¹⁶ I den högra sektionen redovisas samma sex betygsmodeller men nu utifrån en mer fingradig 1–20 skala.¹⁷ Syftet med en mer fingradig betygsskala är att undersöka om en mer fingradig skala i sig själv bidrar till betygens likvärdighet allt annat lika.

Dagens betygssystem (röd stapel) uppvisar en relativt hög grad av olikvärdighet jämfört med de flesta övriga betygsmodeller. Det är bara examensprovsmodellen som uppvisar ett sämre likvärdighetsutfall. Examensprov kan visserligen anses helt rättvisa eftersom alla skriver samma prov och de rättas centralt av oberoende bedömare. Lärarsatta glädjebetyg existerar inte i en sådan modell. Däremot uppstår stora mätfel då elevens hela betyg i ämnet avgörs av prestationen på ett enda prov, vilket förklarar varför examensprovsmodellen uppvisar sämre utfall än dagens betygssystem.

De betygsmodeller som uppvisar lägst olikvärdighet är de som kombinerar ankring av betygen i provresultaten på skolnivå med oberoende nationella slutprov, dvs. att lärarna sätter betyg oberoende av hur eleverna presterar på slutprovet. Dessa två modeller är kombinationsmodellen och modereringsmodellen. Modereringsmodellen klarar sig allra bäst därför att själva kalibreringen sker på ett sätt så att den fördelas rättvist på samtliga elever på skolan. I kombinationsmodellen behöver lärarna själva justera betygen för enskilda elever och är då begränsade av betygsskalans trubbighet. Med modereringsmodellen kan betygen justeras ner med bråkdelar av ett betygssteg för varje elev så att den totala justeringen för hela skolan ändå blir den samma.

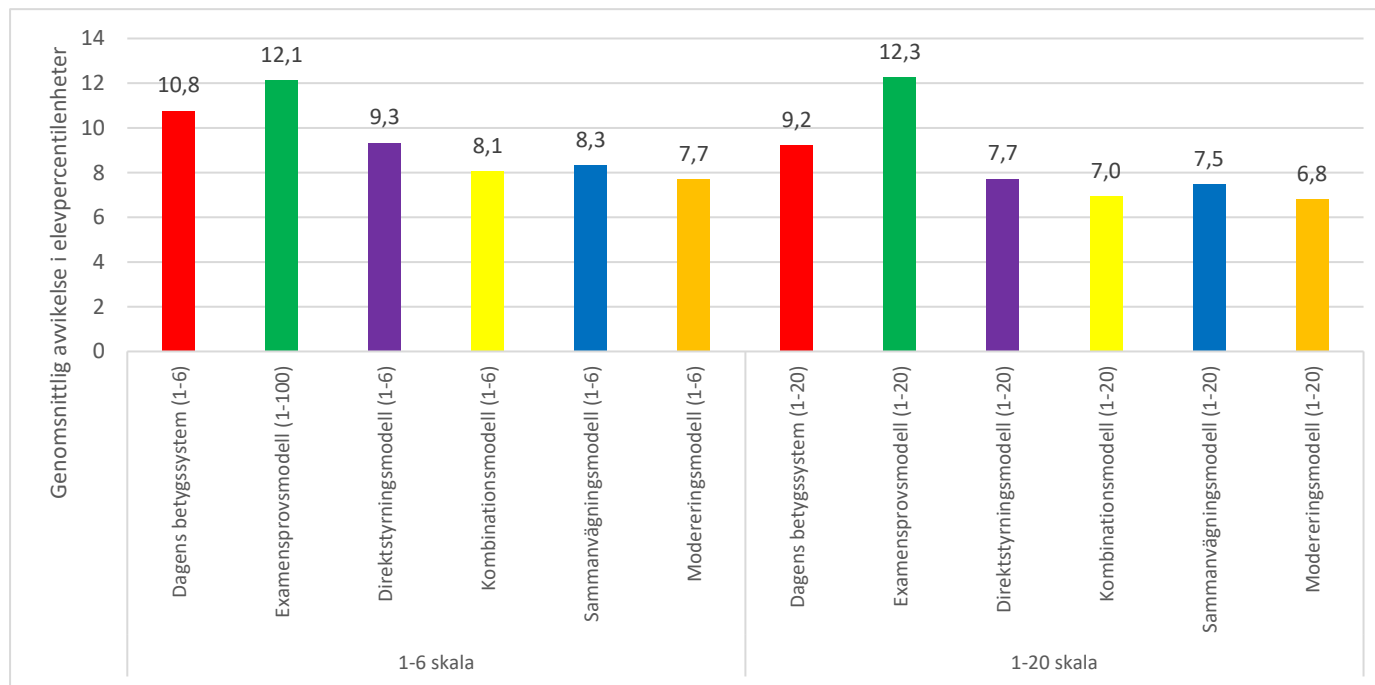
Oavsett vilken betygsskala som används är mönstret detsamma; modereringsmodellen uppvisar lägst olikvärdighet medan examensprovsmodellen uppvisar störst olikvärdighet. Däremot är det värt att notera att oavsett modell så förbättras likvärdigheten med en mer fingradig betygsskala. Detta syns dock inte för examensprovsmodellen eftersom inte en 1–6 skala utan en 1–100-skala använts i simuleringen för denna modell. Hade så varit fallet hade olikvärdig-

¹⁶ Med undantag för examensprovsmodellen där provresultatet på en 1–100 skala utgör betyg.

¹⁷ Här har däremot provresultatet i examensprovsmodellen skalats om till en 1–20 skala. Detta förklarar varför likvärdigheten är sämre för denna modell när alla andra modeller uppvisar förbättringar. Om vi använt en 1–6 skala för examensprovsmodellen i den vänstra delen hade likvärdighetsutfallet varit 13,3 i stället för 12,1 och då hade den 20-gradiga skalan varit bättre även för examensprovsmodellen.

heten varit ännu större för examensprovsmodellen i den vänstra delen av figuren. Det innebär att fler skalsteg i betygsskalan kan bidra till en ökad likvärdighet, eftersom den behåller mer information om elevers kunskaper. Det finns dock andra överväganden angående val av betygsskala som utredningen redovisar i huvudbetänkandets kapitel 9.

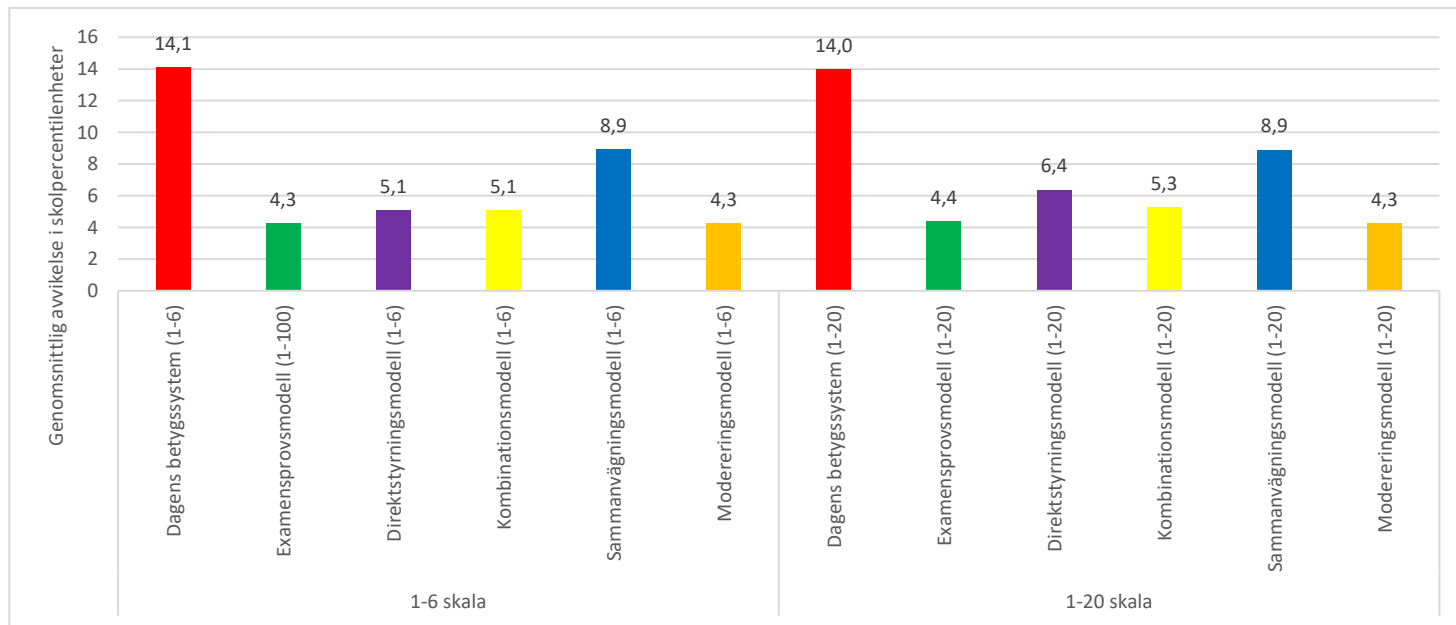
Figur 1.47 Olikvärdighet på elevnivå för olika betygssystem, genomsnitt av de tio ämnen i grundskolan som har nationella prov



Anm.: Högre staplar innebär större olikvärdighet. Måttet är genomsnittlig avvikelse i elevs ranking jämfört med den ranking de skulle fått om de betygssatts perfekt efter deras "sanna kunskap". En perfekt likvärdig betygssystem har värdet 0.

Källa: Utredningens egna beräkningar utifrån Skolverkets statistik.

Figur 1.48 Olikvärdighet på skolnivå för olika betygssystem, genomsnitt av de tio ämnen i grundskolan som har nationella prov



Anm.: Högre staplar innebär större olikvärdighet. Måttet är genomsnittlig avvikelse i elevers ranking jämfört med den ranking de skulle fått om de betygssatts perfekt efter deras "sanna kunskap". En perfekt likvärdig betygssystem har värdet 0.

Källa: Utredningens egna beräkningar utifrån Skolverkets statistik.

I figur 1.48 visas motsvarande olikvärdighet i betygssättning för samma betygsmodeller på skolnivå. Det innebär att vi jämför rangordningen mellan skolor i stället för mellan elever men utifrån samma princip. Skolorna rangordnas efter betygsutfallet enligt en viss betygsmodell och sedan jämförs den rangordningen med den rangordning av skolan som skulle uppstått om skolorna rangordnades efter sina elevers genomsnittliga ”sanna kunskap”.

Skolnivåanalysen visar delvis på annorlunda resultat. Det mest iögonfallande är att examensprovsmodellen från att ha uppvisat minst likvärdighet på elevnivå nu uppvisar lika bra likvärdighet som modereringsmodellen. Detta kan förklaras med att när vi aggregerar provresultaten till skolnivå tar de stora mätfel som uppstår för enskilda elever till stor del ut varandra. Att mäta en skolas ”sanna kunskap” med ett examensprov är med andra ord ett ganska mätsäkert sätt och det är också av den anledningen som vi använder de nationella slutproven för att ankra betygen på skolnivå i direktstyrningsmodellen, kombinationsmodellen samt modereringsmodellen. Att kombinations- och direktstyrningsmodellen ändå inte kommer ner till samma nivå som modereringsmodellen beror på att det fortfarande finns en viss trubbighet i betygen när lärarna tvingas justera betygen för enskilda elever med hela betygssteg i stället för att sprida ut en justering på skolans samtliga elever. Dessutom, och antagligen ännu viktigare, är att varken direktstyrningsmodellen eller kombinationsmodellen, till skillnad från modereringsmodellen, kan kalibrera *spridningen* i betygen utifrån spridningen på provet. Detta leder till att det kan uppstå systematisk olikvärdighet på en skola även om genomsnittet kalibreras. Till exempel om en skola sätter relativt mer generösa betyg för de lågpresterande eleverna än för de högpresterande, då blir spridningen i betyg mindre än vad den borde vara. Detta leder till att även efter att genomsnittet för betygen anpassas till att motsvara genomsnittet för provresultaten så kommer de lågpresterande eleverna att ha fått lite för höga betyg och de högpresterande eleverna lite för låga betyg. Detta får anses vara en klar nackdel med dessa två modeller.

Slutsatsen är att även när vi analyserar likvärdigheten i betygssättning på skolnivå uppvisar modereringsmodellen tillsammans med examensprovsmodellen de bästa resultaten. Observera dock att elevnivån är den mest relevanta nivån eftersom det i slutändan är olikvärdigheten för enskilda elever vi försöker minimera. Att studera olikvärdigheten på skolnivå ger oss dock värdefulla insikter. Vi kan

t.ex. se att sammanvägningsmodellen inte klarar sig lika bra på skolenivå som på elevnivå. Det beror på att lärarsatta betyg inte är ankrade i proven i sammanvägningsmodellen, vilket är en stor svaghet med denna modell. Då har simuleringen inte tagit hänsyn till att det just med sammanvägningsmodellen riskerar att uppstå indirekta effekter, ofta kallade *feedback*-effekter, som gör att skolorna kan komma att vara ännu mer generösa när lärarbetygen inte får mer än ett 50-procentigt genomslag på det slutliga betyget.

Det finns även andra begränsningar i simuleringen som bör nämnas. Simuleringen bygger på att alla prov bedöms centralt och inte av de betygssättande lärarna. Detta gäller även dagens betygssystem så som simuleringen är utförd men är naturligtvis inte helt realistiskt. Därmed har sannolikt olikvärdigheten i dagens betygssystem *underskattats* i denna simulering eftersom det är väl belagt att det finns betydande inslag av olikvärdig bedömning av de nationella proven.¹

1.5.4 Hur kan dessa resultat översättas till en meritvärdeskalibreringsmodell?

Samtliga analyser i avsnitt 1.5 baseras på ämnesvisa simuleringar varpå vi beräknar genomsnittet för samtliga 10 ämnen. Att genomföra analysen utifrån en multivariat ansats har inte varit möjligt inom ramen för denna utredning. Det finns dock mycket som talar för att samma principer som ligger bakom att vissa betygsmodeller klarar sig bättre än andra även gäller på meritvärdesnivå. Dessa principer kan sammanfattas så här:

- Skolans genomsnittsbetyg måste ankra i skolans genomsnittliga provresultat.
- De nationella proven ska vara oberoende mätningar av elevernas kunskaper och lärarna ska därför sätta sina betyg oberoende av elevernas resultat på dessa prov.
- Det slutliga betyget/meritvärdet ska ta vara på den information om elevernas kunskaper som läraren besitter. Därmed ska elevens slutliga betyg utgöras av en kombination av ett kalibrerat lärarbetyg och elevens individuella resultat på slutproven.

¹ Se Tyrefors Hinnerich & Vlachos (2012, 2013).

- En modell där kalibrering görs centralt, ger ett bättre utfall än om enskilda elever får sänkta betyg med hela steg.
- En modell där kalibrering görs centralt ger också större möjligheter att kalibrera inte bara *genomsnittliga* betyg/meritvärden utan även *spridningen* i betyg/meritvärden.

Alla dessa egenskaper uppvisas inte bara av modereringsmodellen utan även av den föreslagna meritvärdeskalibreringsmodellen.

1.6 Skalning av provresultat, betyg och meritvärden

I detta avsnitt behandlas hur gymnasieskolans provresultat från olika provutgåvor kan sättas på samma skala med hjälp av information på gruppnivå från betygen i årskurs 9. Dessutom förs ett resonemang om vilken skala som är lämplig för att återrapportera elevers meritvärde.

1.6.1 Möjlighet att sprida ut elevers provgenomförande på flera år

Om en skolenhet/ett program avslutar den högsta nivån i ett ämne där slutprov ska genomföras före det tredje gymnasieåret, finns möjlighet att genomföra slutprovet redan i slutet av det första eller det andra gymnasieåret. Exempelvis kan det vara aktuellt i matematik (det grundläggande provet i vårt förslag) där många skolor/program i nuvarande system genomför kurserna matematik 1a/1b/1c under det första läsåret på gymnasieskolan. Det är dock nödvändigt att samtliga elever på programmet på skolenheten skriver proven samtidigt eftersom de kalibreras på denna nivå mot samma provutgåvor. Därför kan inte elever inom olika programinriktningar eller enskilda elever skriva proven olika år för att de läser ämnena på fler nivåer än sina programkamrater. En skolas val kan medföra fördelar eller nackdelar för olika elever. Exempelvis är det i ämnet samhällskunskap på samhällsprogrammet endast obligatoriskt att läsa nivå 1 för programinriktningen medier, information och kommunikation. Inriktningen beteendevetenskap måste däremot även läsa nivå 2 och inriktningen samhällsvetenskap läser till nivå 3. En skola som erbjuder samtliga dessa inriktningar antas då välja att genomföra slutprovet år 3. Det

kan vara till nackdel för eleverna som avslutade ämnet tidigare och har tappat kunskaper. Ett annat alternativ är att genomföra provet tidigare, vilket blir till nackdel för de elever som ska läsa vidare och skulle kunna prestera bättre på provet i slutet av år 3. Skolan kan givetvis införa ett obligatorium för samtliga samhällselever att välja fler kurser i samhällskunskap så det blir lika för alla eller erbjuda inriktningar inom program där problematiken inte uppstår.

Att vissa skolor/program genomför slutprovet i grundläggande matematik i slutet av år 1, andra i slutet av år 2 och ytterligare andra i slutet av år 3 innebär att eleverna inte kommer att genomföra samma provutgåva. Det skulle kunna innebära en orättvisa om elever som går ut gymnasieskolan samma år har genomfört olika provutgåvor om den årsvisa standardiseringen baseras på olika elevsammansättningar. En sådan situation skulle kunna uppstå om den elevgrupp som skriver t.ex. grundläggande prov i matematik efter första året i gymnasieskolan skiljer sig från den elevgrupp som skriver motsvarande prov efter andra eller tredje året i gymnasieskolan.

Detta är inte nödvändigtvis ett problem eftersom flexibiliteten för skolor innebär att proven för varje enskilt år kommer att genomföras av elever från samtliga tre gymnasieår. Så länge proportionerna mellan dessa är någorlunda lika över tid kommer den årsvisa standardiseringen att göra proven jämförbara eftersom elevsammansättningen varje år kommer att vara lika. I figur 1.49 visas en schematisk bild över hur det kan se ut.

Figur 1.49 Hypotetiskt exempel på hur olika skolor kan välja att genomföra det grundläggande matematik-provet under olika gymnasieår

Årsvis standardiserade prov (jämförbara prov)		2028	2029	2030	2031	2032
Avgångskull 2030		År 1	År 2	År 3		
	Skola A	mat1				
	Skola B		mat1			
	Skola C			mat1		
Avgångskull 2031			År 1	År 2	År 3	
	Skola A		mat1			
	Skola B			mat1		
	Skola C				mat1	
Avgångskull 2032				År 1	År 2	År 3
	Skola A			mat1		
	Skola B				mat1	
	Skola C					mat1
Avgångskull 2033					År 1	År 2
	Skola A				mat1	
	Skola B					mat1
	Skola C					
Avgångskull 2034						År 1
	Skola A					mat1
	Skola B					
	Skola C					
Antal prov från olika skoltyper		1	2	3	3	3

Anm.: Utredningens visualisering.

I exemplet antas att proven börjar genomföras våren 2028 så att den årskull som påbörjat gymnasieskolan hösten 2027 och går ut våren 2029 kan skriva matematik-grundprovet redan 2028. I exemplet finns endast tre skolor men som var och en kan tänkas representera ett stort antal skolor som väljer att genomföra provet vid samma tillfälle. Skola A genomför provet i slutet av år 1 varje år, skola B genomför det år 2 varje år och skola C genomför det år 3 varje år.

Antag att skola A är en relativt högpresterande skola, skola B en genomsnittlig skola och skola C en lågpresterande skola. Detta innebär sammantaget att år 2028, när provsystemet införs, är det endast skola A som skriver det grundläggande matematikprovet. Om vi årsvis standardiserar provet, givet att de elever som genomförde provet i genomsnitt var högpresterande, kommer dessa elever att missgynnas av den årsvisa standardiseringen, relativt elever från samma årskull men som skriver provet i år 2 eller år 3. Även år 2029 uppstår rättviseproblem om provet standardiseras årsvis eftersom nu visserligen elever i skola B skriver provet samt nästa årskull från skola A, men

eleverna i skola A och B är inte representativa för hela årskullen eftersom inga lågpresterande elever (skola C) är med. Vi ska återkomma till problemet med att standardisera proven för de inledande åren senare.

När vi kommer till 2030 skriver däremot elever från samtliga tre skolor det grundläggande matematikprovet. De tillhör visserligen inte samma årskull men det spelar ingen roll så länge varje årskulls genomsnittliga prestationsnivå är någorlunda lika (vilket de med största sannolikhet är eftersom förändringar i genomsnittlig prestationsnivå för intelligande årskullar är mycket små, särskilt i förhållande till förändringar i provens svårighetsgrad). Detta gäller också stora undergrupper av en hel årskohort, t.ex. samtliga skolor som väljer att skriva provet ett visst år och som i detta exempel är representerade av skola A, B och C.² Det spelar heller ingen roll om skolorna är olika stora till elevantal. Åren 2030, 2031 och 2032 kommer sammansättningen av elever som genomför respektive prov att vara representativ för en enskild årskull, trots att eleverna kommer från tre olika årskullar.

Detta förstås enklast om man studerar de tre vertikala grå segmenten i figur 1.49. Utifrån de antaganden som gjorts kommer elevsammansättningen för de grundläggande matematikproven att vara identisk för varje år 2030, 2031 samt 2032 (och efterföljande år) och den årsvisa standardiseringen gör då att provresultaten blir jämförbara mellan åren. Därmed spelar det ingen roll att eleverna i de olika skolorna skrivit olika prov när betygsvärdena ska kalibreras efter år 3, baserat på ett samlat provresultat från ytterligare provämnerna.

För åren 2028 och 2029, som är en övergångsperiod i samband med införandet av de nya proven, är däremot elevsammansättningen, som vi redan konstaterat, inte representativ för hela årskohorten och kan därför inte standardiseras på samma sätt. Vi har antagit att skola A är högpresterande och skola B presterar på en genomsnittlig nivå. Därmed kommer de elever som skriver provet 2028 respektive 2029 i genomsnitt prestera på en högre nivå än genomsnittet för hela årskullen (elever från alla skolor A, B och C) och hänsyn måste därför tas till detta när dessa prov ska standardiseras (placeras på samma skala).

² Observera att för en enskild skola är det naturligtvis inte realistiskt att anta att skolans prestationsnivå är konstant från ett år till nästa, men i verkligheten representerar varje skola A, B respektive C tillsammans alla skolor i hela riket men med olika prestationsnivå. Då är det betydligt mer realistiskt att anta t.ex. att prestationsnivån för hundratals typ A-skolor är konstant mellan intelligande år.

Ett sätt att göra detta är att utnyttja informationen från betygen eller slutprovsresultaten i matematik från årskurs 9 i grundskolan. En invändning skulle kunna vara att dessa betyg inte är pålitliga om de inte kalibrerats och vi kalibrerar inte enskilda betyg i meritvärdeskalibreringsmodellen. I verkligheten representerar skola A respektive B i exemplet egentligen grupper av skolor och vi är endast intresserade av hur de eleverna på dessa skolor presterar som grupp. För 2028 kan vi därför beräkna genomsnittet i betyg eller provresultat ifrån grundskolan för samtliga elever i skola A och jämföra med motsvarande genomsnitt för hela elevkohorten. Antag att det genomsnittliga resultatet i matematik från slutprovet i årskurs 9 för eleverna i gymnasieskola A är 10 standardiserade poäng högre än motsvarande standardiserade riksgenomsnitt. Då placerar vi genomsnittet för det grundläggande matematikprovet 10 standardiserade poäng högre än de standardiserade genomsnitten för motsvarande prov 2030 (och där genomsnittet för 2030 är det samma som för 2031 och 2032).

För 2029 använder vi samma princip men beräknar då genomsnittet baserat på både elever i skola A och skola B eftersom det är dessa elever som skrivit det grundläggande matematikprovet 2029.

Att utnyttja betyg och eller provresultat från årskurs 9 är inte bara praktiskt, utan kan också vara ett lämpligt kvalitetsgranskningsverktyg för att ytterligare säkra den årsvisa standardiseringen för proven 2030, 2031 och 2032. Detta för att se om det uppstår stora förändringar i vilka skolor som väljer att göra proven från år till år och att därför elevsammansättningen inte blir tillräckligt stabil över tid för dem som skriver provet.

För de prov som enbart genomförs sista året på gymnasieskolan av samtliga skolor, t.ex. avancerad matematik är inte detta nödvändigt eftersom alla elever skriver samma provutgåva.

Sammanfattningsvis är det alltså inte ett problem om olika skolor och program tillåts att genomföra vissa av slutproven i slutet av år 1 eller 2 i stället för i slutet av år 3. Det kan däremot diskuteras om alla elever på samma program och skola måste genomföra provet vid samma tillfälle eller om elever på olika programinriktningar inom samma skola ska kunna genomföra prov vid olika tillfällen. Teoretiskt bör det vara möjligt att även tillåta olika provår inom ett program på samma skola om vi definierar en programinriktning som ett eget program och det finns tillräckligt många elever på inriktningen. Om provresultaten från olika provutgåvor standardiseras på ett sådant

sätt att de är jämförbara ska det inte spela någon roll om elever som tillhör samma program på samma skola har skrivit olika prov eftersom deras provresultat är jämförbara. En viss ökning av mätosäkerheten kommer däremot att uppstå och det kan då finnas skäl att komplettera den årsvisa standardiseringen med justeringar baserat på hur elevsammansättningen varierar för olika provutgåvor, baserat på betyg eller provresultat från årskurs 9.

1.7 Vilken skala/skalor är lämpligast för återrapportering av betyg, betygsvärde, provresultat respektive meritvärde?

1.7.1 Provskala

I kapitel 17 gav vi förslaget att de nationella proven ska ha en nyanserad skala som inte bjuder in till tolkning i termer av betyg förrän provresultatet räknas samman. Olika provutgåvor kan dessutom ha olika antal maximala poäng eftersom dessa kommer att skalas om till en gemensam skala för olika utgåvor och prov så att de blir jämförbara. Vi föreslår t.ex. en 1–100 skala. Det går inte att dividera provpoängen 1–100 med 10 för att översätta till betygen 1–10 eftersom provpoängen bygger på medelvärde och fördelning.

Provresultaten måste någon gång översättas så att de blir jämförbara med betygens 1–10 skala. Om denna skala bara kan anta heltalsvärden blir skalan grövre än provens poäng, vilket innebär en informationsförlust när provpoängen omvandlas. Vi föreslår därför att provresultatet redovisas med decimaler. Det innebär ändå viss informationsförlust. Det ska därmed göras så sent som möjligt i processen. Proven ska ha en betydligt mer nyanserad skala för såväl poäng från uppgifter (råpoäng) som poäng efter utgåvans omskalning (skalpoäng) i ekvivaleringen. Det är önskvärt att undvika intervall för olika poäng t.ex. att 59 till 67 poäng skall omvandlas till 7. När provpoäng översätts till skalan 1–10 kommer det då alltid finnas elever som har fått enstaka poäng under eller över gränsen för nivåerna 1–10 och de som överskrider gränsen med råge. De elever vars poäng ligger mitt emellan två gränser kan sägas med större säkerhet ha fått korrekt provresultat än de som ligger nära gränsen. Det är därför bättre att skala om elevernas provpoäng t.ex. 64 poäng till ett värde med decimaler t.ex.

7,32. För en elev som inte känner till reglerna för omskalning av provpoäng till provresultat säger dock inte 64 poäng av 100 mer än 7,32 av 10 och provpoängen behövs aldrig redovisas för eleven.

I ämnena engelska, svenska och svenska som andraspråk har olika delprov behandlat olika färdigheter (lyssna, läsa, skriva) i språk. Elevernas lärare har använt delprovresultat för att jämföra sitt underlag med resultatet på proven med avseende på dessa färdigheter. Eftersom eleverna har fått betydligt färre chanser att visa vad de kan med avseende på enskilda färdigheter än de fått att visa sina kunskaper i ämnet som helhet är informationen från ett delprov mindre säker än resultatet på hela provet. Tanken är inte heller att lärarna ska hitta förklaringar till elevers provresultat efter betygssättningen. Eftersom lärarna inte får tillgång till enskilda elevers resultat är det inte meningsfullt att återkoppla resultat på delprov. En elev kan dock själv granska bedömningen av olika uppgifter i sitt eget prov.

1.7.2 Meritvärdesskala

Såväl betygsvärden som meritvärden behöver redovisas i decimalform. Det är nödvändigt för att elevernas meritvärden ska spridas ut mer inför antagning till vidare studier. Man kan då tänka sig att man översätter denna 1–10-skala till en heltalsskala med fler skalsteg. En sådan översättning innebär att kopplingen mellan betyg och meritvärden blir otydlig. Om eleven fått betygsvärdet 6,7 provresultatet 65 och meritvärdet 45 blir det svårt att se vad som hänt under vägen. Fördelen med att ha olika skalor i betyg, prov respektive meritvärden är att de tre värdena får sina egna betydelser som inte direkt ska översättas till de andra och därmed inte öppna upp för osäkra tolkningar. Nackdelen är att beräkningarna dem emellan sker i en svart låda där elever m.fl. bara får acceptera utfallet. Vi förordar därför att såväl provresultat som meritvärden återkopplas i en 1–10-skala med två decimalers noggrannhet.

1.8 Korrelationer (samband) mellan resultat i ämnen med nationella prov och ämnen utan nationella prov

En hypotes som har undersökts är om betyg i ämnen i vilka det inte finns nationella prov skulle kunna kalibreras med ämnen med nationella prov. Vi har därför utvärderat hur det skulle fungera att använda resultaten från ett nationellt prov i ett ämne för att kalibrera betyg på gruppnivå (skolenhetsnivå) i ett annat ämne. Ämnesgrupperna som har undersökts är moderna språk, NO och teknik, SO och PREST.

1.8.1 Data och metod

Korrelationsanalyserna baseras på uppgifter för åren 2015–2023. De sista observationerna gäller de som slutade årskurs 9 vårterminen 2023. Åren 2020–21 genomfördes inga nationella prov på grund av pandemin så för dessa år finns inga data. Datasetet utgörs av data över slutbetyg i årskurs 9 i grundskolan och nationella prov i årskurs 9 i grundskolan.

Moderna språk som elevens val är tämligen ovanligt. Åren 2015–2023 var det enbart 9 816 betyg registrerade (år 2023 var det 945 betyg registrerade), därför tas inte moderna språk som individuellt val med i redovisningarna. Det vanligaste är att eleverna läser moderna språk som språkval i framför allt tyska, franska och spanska. År 2023 hade 84 300 betyg i moderna språk som språkval registrerats. Det är framför allt moderna språk som språkval som är relevant för utredningen.

Korrelationerna är beräknade årsvis för åren 2015–2023. Resultatet är även redovisade som ett genomsnitt för dessa år. Syftet med den sistnämnda metoden är att eliminera att de nationella proven tenderar att variera i svårighetsgrad. Vi lämnar förslag på hur proven kan stabiliseras, men för att inte bygga in provens varierande svårighetsgrad i våra analyser undersöker vi medelvärdet i korrelationerna för flera år så att systematiska skillnader i provens svårighetsgrad inte är det som gör att korrelationerna blir starka eller svaga. Det kan även finnas andra orsaker till att korrelationerna varierar över tid som inte har med provens svårighetsgrad att göra, t.ex. variation i vilka domäner som täcks.

Syftet med undersökningen är att ta reda på om det går att kalibrera t.ex. moderna språk med engelska på gruppnivå, dvs. inom en skolenhet. Därför undersöks även korrelationerna inom skolenheterna/skolenhetsvis. För att undersöka hur korrelationen mellan provbetyg och betyg påverkas även på skolenhetsnivå aggregeras resultaten till skolenhetsnivå. Beräkningen är gjord baserat på uppgifter från elever som både har betyg i ämnet och resultat på det nationella provet som sambandet undersöks med. I och med att skolorna varierar i storlek viktas de utifrån storlek i korrelationsberäkningen.

Regressionsmodeller har även använts för att pröva hypotesen att flera ämnen med nationella prov sammantaget kan kalibrera betyg i t.ex. PREST-ämnena. I modellerna används skolenhet som oberoende variabel. Ett exempel på en sådan modell är när vi undersöker om flera ämnen med nationella prov tillsammans skulle kunna kalibrera nivån för betyg i PREST-ämnena genom att vardera PREST-ämne är oberoende variabel med det nationella provet i matematik, svenska och engelska som kontrollvariabler.

Vi har även undersökt korrelationer mellan provresultaten i sin helhet och betygsvärden (betygen beräknade på en skala 1–6) respektive meritvärdet (jämförelsetalet) i sin helhet på skolenhetsnivå. Detta beräknas utifrån samma dataunderlag som använts för beräkning av meritvärdeskalibrering för grundskolan respektive gymnasieskolan.

1.8.2 Resultat och slutsatser

Resultaten indikerar att det nationella provet i ett av NO- respektive SO-ämnena kan användas för att kalibrera övriga betyg i NO- respektive SO-ämnena på skolenhetsnivå. Teknikämnet har något sämre korrelationer än NO-gruppen, men även där bedöms det rimligt att kalibrera med ett nationellt prov i NO. Ytterligare analyser har gjorts såsom kontroll för skolenhetsnivå för NO och SO, men skillnaderna var små.

Betygen i moderna språk som språkval visar däremot endast ett svagt samband med det nationella provet i engelska. Det bedöms därför tveksamt om det kan anses ändamålsenligt att kalibrera betyg i moderna språk med det nationella provet i engelska.

Korrelationerna mellan ämnena med nationella prov och betygen i PREST-ämnena är låga på både individ- och skolenhetsnivå.

Inte heller en mer komplex modell, där det sammanvägda värdet utifrån de nationella proven i matematik, svenska och engelska nyttjas, tycks ge särskilt stor effekt. Modellernas sammantagna förklaringsvärde ligger på ett R^2 runt 0,3. I två av modellerna läggs nettoavvikelserna för matematik, svenska och engelska till. Detta förbättrar modellen något, men det är fortfarande låga förklaringsgrader (och likaså korrelationer).

Nettoavvikelserna visar dock på ett annat intressant samband, nämligen att stora avvikelser mellan lärarsatta betyg och nationella provresultat på en skola korrelerar med högre betyg i PREST-ämnena. Det tolkar vi som att en generös lärarbetygssättning i de nationella provämnena även återspeglas i en generös betygssättning i PREST-ämnena.

Korrelationen mellan resultatet på samtliga nationella prov och betygsvärdet respektive meritvärdet på skolenhetsnivå ligger högt. För grundskolan ligger korrelationen mellan betygsvärdet och nationella provresultatet på skolenhetsnivå på 0,90 och 0,87 för 2022 respektive 2023. För gymnasieskolan ligger det på 0,82 och 0,84 för 2018 respektive 2019.

1.9 Alternativa kalibreringsmodeller – Ekvipercentilkalibrering

I detta avsnitt redovisas kortfattat hur meritvärdeskalibrering kan tillämpas utifrån en algoritm som bygger på ekvipercentilkalibrering i stället för den linjära modell som redovisats i avsnitt 1.1 och som ligger till grund för alla resultat som redovisats hittills i denna statistiska bilaga.

Utredningen avgränsar sig ifrån att fullt ut utvärdera exakt vilken typ av kalibreringsalgoritm som ska tillämpas, detta måste utredas vidare av lämplig myndighet men bör inte vara en fråga för politiken utan snarare en teknisk fråga för experter på lämplig myndighet.

Det finns i skrivande stund minst tre olika algoritmiska modeller som potentiellt kan användas för meritvärdekalibrering enligt de koncept som utredningen föreslår:

- Linjär modell
- Kvadratisk modell
- Kvartilmodell

I detta avsnitt beskrivs hur kvartilmodellen fungerar via hypotetiska exempel och där utfallet kan jämföras med den linjära modellen.

1.9.1 De grundläggande skillnaderna mellan kvartil- och linjär modell

I den linjära modellen styrs själva kalibreringen av två parametrar som beräknas utifrån den hypotetiska Elefantskolans resultat på de nationella proven:

1. Skolans genomsnittliga provresultat
2. Skolans standardavvikelse i provresultat (spridningen)

Utifrån dessa parametrar kalibreras varje elevs betygsvärde, vilket visades i figur 1.1–1.3 i avsnitt 1.1.

I en kvartilmodell (även kallad ekvippercentilmodell) kalibreras en skolas elever i stället utifrån fyra parametrar som beräknas utifrån skolans resultat på de nationella proven:

Q4: 4:e kvartilen = det högsta provresultat,

Q3: 3:e kvartilen (= 75:e percentilen),

Q2: 2:a kvartilen = medianen (= 50:e percentilen),

Q1: 1:a kvartilen (= 25:e percentilen).

Det högsta provresultatet är lika högt som högre än samtliga provresultat på skolan och motsvarar den 4:e kvartilen eller den 100:e percentilen. Den tredje kvartilen är det provresultat som är lika högt eller högre än 75 % av alla provresultat. Den 2:a kvartilen är det värde som är lika högt eller högre än 50 % av alla provresultat på skolan och ligger således i mitten av provbetygsfördelningen och kallas även för medianvärdet. Slutligen är den 1:a kvartilen det värde som är minst lika högt som 25 % av alla provresultat.

1.9.2 Kvartilmodellen steg för steg

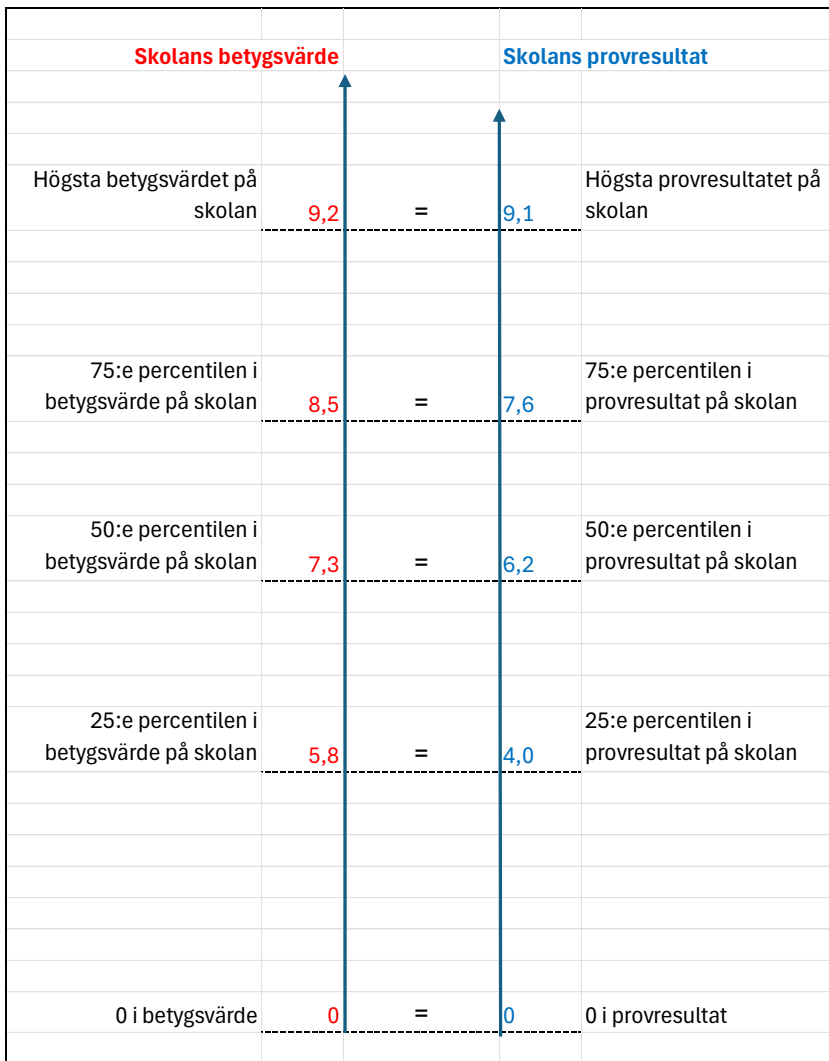
I ett första steg placeras en skolas lärarsatta betygsvärden på samma skala som Elefantaskolans samlade provresultat (där provresultaten standardiserats på nationell nivå till 1–10 skala med medelvärde 6,0 och standardavvikelse 1,5). Detta går till enligt följande och åskådliggörs i figur 1.50:

1. Det högsta betygsvärdet ekvivaleras med det högsta provresultatet (som per definition också är den 4:e kvartilen).
2. Det betygsvärde som motsvarar den 3:e kvartilen ekvivaleras med motsvarande 3:e kvartil i provresultat.
3. Det betygsvärde som motsvarar den 2:a kvartilen (= medianvärdet) i betygsvärde ekvivaleras med motsvarande 2:a kvartilvärde i provresultat.
4. Det betygsvärde som motsvarar den 1:a kvartilen ekvivaleras med motsvarande 1:a kvartilen i provresultat.

Det första steget innebär att vi har fixerat fyra punkter på betygsvärdesskalan som ekvivaleras med motsvarande punkter på provresultatsskalan. Observera att vi ännu inte kalibrerat någon elevs betygsvärde, vi har endast satt betygsvärden och provresultat på en gemensam skala. Se figur 1.50 för en illustration av detta steg.

En viktig sak att observera är att så som de två skalorna är framställda i figuren är de inte nödvändigtvis proportionerliga mot varandra. Till exempel kan ses att avståndet mellan den 3:e och 4:e kvartilen är 0,7 poäng på betygsvärdesskalan respektive 1,5 poäng på provresultatsskalan. Det innebär att i detta specifika intervall motsvarar ett avstånd på 0,7 poäng i betygsvärde 1,5 poäng i provresultat. Men denna proportion gäller inte (nödvändigtvis) för övriga intervall på skalorna. För intervallet mellan medianvärdet eller 2:a kvartilen och den 3:e kvartilen motsvarar 1,2 poäng på betygsvärdesskalan 1,4 poäng på provresultatsskalan i detta exempel. I botten på skalan ser vi att 5,8 betygsvärdespoäng motsvarar 4,0 provresultatspoäng. Figuren är således inte skalenlig men visar hur flexibel kvartilmodellen är då den fångar upp att relationen mellan betygsvärde och provresultat kan vara olika beroende på var i betygsvärdesfördelningen en elev är.

Figur 1.50 Steg 1 i Kvartilmodellen: Ekvivalering av skalan för skolans betygsvärden och skalan för skolans provresultat



Anm.: Utredningens visualisering.

I nästa steg kalibreras enskilda elevers betygsvärden till kalibrerade betygsvärden, vilket åskådliggörs i figur 1.51.

Givet att vi ankrat skalan för betygsvärdena i skalan för provresultat med hjälp av de fyra kvartilvärdena, kan vi nu placera varje elevs betygsvärde på betygsvärdesskalan (den vänstra) och läsa av

horisontellt det kalibrerade betygsvärdet på den högra skalan (som är identisk med provresultatsskalan).

Till exempel kommer Kristina, som hade det högsta betygsvärdet på skolan att få ett kalibrerat betygsvärde som motsvarar det högsta provresultatet, vilket innebär att Kristinas betygsvärde på 9,2 poäng kalibreras till 9,1 poäng.

Maria som har 8,6 i betygsvärde får ett kalibrerat betygsvärde på 7,8. För att förstå varför hennes betygsvärde kalibrerades ner mer än Kristinas (0,8 poäng mot 0,1 poäng), låt oss gå igenom den algoritm som styr kalibreringen:

Marias betygsvärde ligger i intervallet mellan 3:e och 4:e kvartilen men hon ligger betydligt närmare den 3:e kvartilen än den 4:e, närmare bestämt $(8,6-8,5)/(9,2-8,5) = 0,14$, dvs. Marias betygsvärde ligger 14 % in i intervallet mellan Q3 och Q4. Eftersom intervallets längd är 0,7 poäng på betygsvärdesskalan motsvarar det $0,14 \cdot 0,7 = 0,098 \approx 0,1$ poäng, vilket vi kan läsa av direkt i figur 1.50. Detta innebär att Marias kalibrerade betygsvärde också ska ligga 14 % in i intervallet på provresultatsskalan, vilken per definition är identisk med den kalibrerade betygsvärdesskalan. Marias kalibrerade betygsvärde ska således ligga $0,14 \cdot (Q4 - Q3) = 0,14 \cdot (9,1 - 7,6) = 0,2$ poäng in på den kalibrerade betygsvärdesskalan. Marias kalibrerade betygsvärde beräknas då slutligen till $0,2 + Q3 = 0,2 + 7,6 = 7,8$ poäng. Marias betygsvärde på 8,6 poäng motsvarar således ett kalibrerat betygsvärde på 7,8 poäng.

Men vi har fortfarande inte besvarat varför Maria kalibrerades ner mer än Kristina. Orsaken är att de lärarsatta betygen (betygsvärdena) var mer sammanpressade än vad provresultaten indikerade i detta intervall. Kvartilekvivaleringen anpassar spridningen i betygsvärden så att den motsvarar spridningen i provresultat. I den linjära modellen görs detta för alla elever på skolan utifrån standardavvikelsen, dvs. om spridningen i hela skolans betygsvärden t.ex. är mindre än spridningen i hela skolans provresultat, då sprids elevernas kalibrerade betygsvärden ut så att standardavvikelsen i kalibrerade betygsvärden motsvarar standardavvikelsen i provresultat. I kvartilmodellen används inte standardavvikelsen för skolan utan i stället anpassas varje kvartilintervall, dvs. avståndet mellan Q4 och Q3, Q3 och Q2 osv. utifrån att vi fixerar dessa kvartilvärden på respektive skala. Fördelen med detta är att kalibreringen är mer flexibel så att

1.9.3 Kvartilmodellen vs den linjära modellen – ett exempel

I tabell 1.9 och 1.10 visas de data som ligger till grund för exemplet med Elefantskolan och kvartilkalibreringen. I tabellerna visas även utfallet i för den linjära modellen.

Färgsegmenten representerar respektive kvartilintervall för betygsvärde respektive nationella provresultat. Observera att eleverna är rangordnade efter betygsvärdet (kolumn 2). Det innebär att elevernas provresultat inte är rangordnade (kolumn 3) men visas i rangordnad form i kolumn 6. Provresultaten i kolumn 6 hör därför inte nödvändigtvis ihop radvis (elevvis) med övriga kolumner.

I de tre sista kolumnerna jämförs de två kalibreringsmodellerna först med de lärarsatta betygsvärdena (kolumn 9 respektive 10), sedan med varandra (kolumn 11). De grönmarkerade värdena visar att eleven tjänade på kvartilkalibreringsmodellen medan de rödfärgade värdena indikerar att eleven tjänade på den linjära modellen. Det finns en tydlig systematik i vilka som vinner och förlorar på en kvartilmodell jämfört med en linjär modell i detta exempel. De mest högpresterande eleverna och de mest lågpresterande eleverna tjänar på kvartilmodellen även. Detta ska inte ses som ett generellt utfall utan beror på hur detta hypotetiska exempel är konstruerat. Tolkningen är att även om skolan generellt har satt relativt generösa betyg (det genomsnittliga betygsvärdet samt medianvärdet är högre än motsvarande genomsnittliga provresultat respektive medianprovresultat), har de mest högpresterande eleverna (enligt betygsvärdet) inte blivit fullt lika generöst betygsatta som eleverna i mitten av fördelningen. Samtidigt indikerar provresultaten att de allra lägst presterande eleverna blivit relativt sett hårdare betygsatta än eleverna i mitten av fördelningen och att de kompenseras något för det med kvartilmodellen.

Med den linjära modellen hade de mest högpresterande eleverna kalibrerats ner mer eftersom den linjära modellen förutsätter att betygssättningen är mer eller mindre proportionell, såvida den inte är systematisk på ett sätt som går att mäta och därmed även styra med standardavvikelsen för hela skolan.

Tabell 1.9 Linjär respektive kvartilmodell för Elefantskolan

			Linjär modell		Kvartilmodell			Kalibrerade BV jämfört med lärsatta BV		Jämförelse av modeller
Elevrankning	Betygsvärde (BV)	Resultat nationella slutprov (elev)	Kalibrerat betygsvärde (BVK)	Meritvärde* (MV)	Resultat nationella slutprov (rangordnat)	Kalibrerat betygsvärde (BVK)	Meritvärde* (MV)	Diff (BVK_lin-BV)	Diff (BVK_qrt-BV)	Diff (BVK_qrt-BVK_lin)
Kristina 1	9,2	9,1	8,7	8,8	9,1	9,1	9,1	-0,52	-0,10	0,42
2	9,1	8,6	8,6	8,6	9,0	8,9	8,8	-0,55	-0,21	0,33
3	9,1	9,0	8,6	8,7	8,6	8,9	8,9	-0,55	-0,21	0,33
4	8,9	8,2	8,3	8,3	8,2	8,5	8,4	-0,60	-0,44	0,16
Maria 5	8,6	7,6	7,9	7,8	7,7	7,8	7,8	-0,68	-0,79	-0,11
6	8,5	7,7	7,8	7,8	7,6	7,6	7,6	-0,70	-0,90	-0,20
7	8,5	7,5	7,8	7,7	7,6	7,6	7,6	-0,70	-0,90	-0,20
8	8,2	7,6	7,4	7,5	7,5	7,2	7,3	-0,78	-0,96	-0,18
Jens 9	7,9	7,1	7,0	7,1	7,1	6,9	6,9	-0,86	-1,03	-0,16
10	7,7	6,4	6,8	6,7	6,7	6,6	6,6	-0,92	-1,07	-0,15
11	7,6	5,9	6,7	6,4	6,4	6,5	6,3	-0,94	-1,09	-0,15
12	7,4	6,2	6,4	6,3	6,2	6,3	6,2	-1,00	-1,13	-0,13
13	7,2	6,1	6,2	6,1	6,1	6,0	6,0	-1,05	-1,19	-0,15
14	6,9	5,9	5,8	5,8	5,9	5,6	5,7	-1,13	-1,32	-0,20

Hannes 15	6,7	6,7	5,5	5,9	5,9	5,3	5,7	-1,18	-1,41	-0,23
16	6,1	5,3	4,8	4,9	5,3	4,4	4,7	-1,34	-1,67	-0,33
17	6,0	3,9	4,6	4,4	4,1	4,3	4,2	-1,37	-1,71	-0,35
18	5,8	4,0	4,4	4,3	4,0	4,0	4,0	-1,42	-1,80	-0,38
19	5,8	4,0	4,4	4,3	4,0	4,0	4,0	-1,42	-1,80	-0,38
Anders 20	5,4	3,7	3,9	3,8	3,9	3,7	3,7	-1,52	-1,68	-0,15
21	5,4	4,1	3,9	3,9	3,7	3,7	3,8	-1,52	-1,68	-0,15
22	5,1	3,5	3,5	3,5	3,5	3,5	3,5	-1,60	-1,58	0,02
23	5,0	2,8	3,4	3,2	2,8	3,4	3,3	-1,63	-1,55	0,08
24	3,5	2,7	1,5	1,8	2,7	2,4	2,5	-2,03	-1,09	0,94
Medelvärde	7,1	5,98	5,98	5,98	5,98	5,93	5,95	-1,08	-1,14	-0,05
Standardavvikelse	1,55	1,95	1,95	1,94	1,95	1,97	1,96			
Medianvärde	7,30	6,15	6,28	6,24	6,15	6,14	6,14			

Anm.: De rangordnade provresultaten i kolumn 6 hör inte ihop radvis med övriga observationer. Elevens eget provresultat syns i kolumn 3. Meritvärdet för såväl linjär- som kvartilmodell beräknas som ett viktat genomsnitt av det kalibrerade betygsvärdet (70 %) och elevens individuella provresultat (30 %).

Utredningens visualisering.

Tabell 1.10**Kvartilvärden**

Max betygsvärde Q4	9,2	Max nationella slutprov Q4	9,1
Betygsvärde Q3	8,5	Nationella slutprov Q3	7,6
Betygsvärde Q2	7,3	Nationella slutprov Q2	6,2
Betygsvärde Q1	5,8	Nationella slutprov Q1	4,0
Betygsvärde 0	0,0	Nationella slutprov 0	0,0

Längd på respektive kvartilintervall

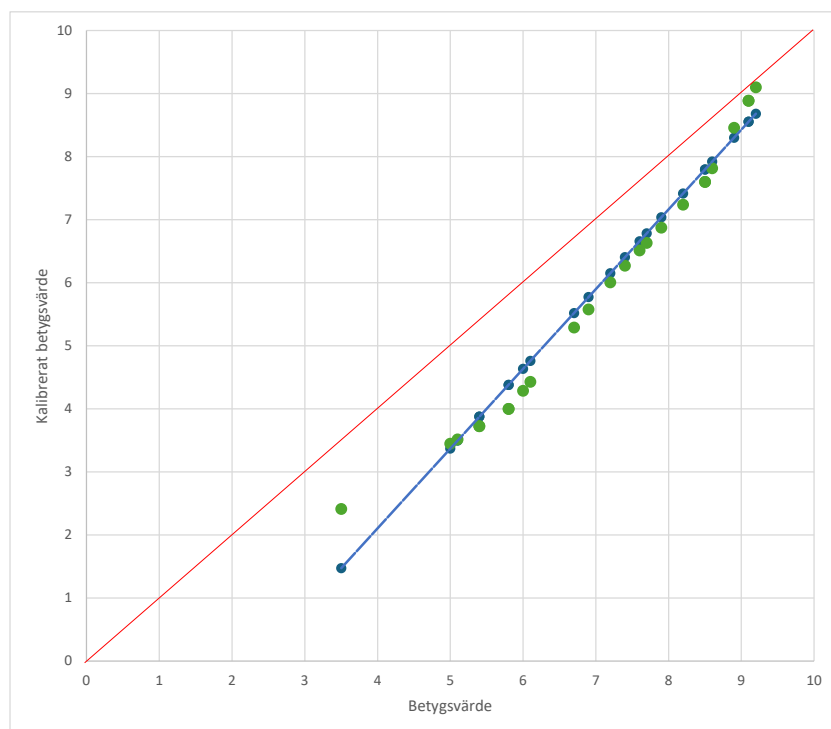
Max betygsvärde Q4	0,7	Max nationella slutprov Q4	1,5
Betygsvärde Q3	1,2	Nationella slutprov Q3	1,5
Betygsvärde Q2	1,5	Nationella slutprov Q2	2,2
Betygsvärde Q1	5,8	Nationella slutprov Q1	4,0

Anm.: Utredningens visualisering.

I figur 1.53 visas slutligen hur de olika kalibreringsmodellerna slår, baserat på samma data som presenterats i tabell 1.9.

I figur 1.53 visas på den horisontella axeln betygsvärdena för eleverna och på den vertikala axeln de kalibrerade betygsvärdena där varje elev representeras med en punkt, blå för den linjära modellen och grön för kvartilmodellen. Eftersom det är två kalibreringsmodeller innebär det att det i princip finns två punkter per elev, men de är på samma plats på x-axeln eftersom de har samma lärarsatta betygsvärde, oberoende av kalibreringsmodell. På y-axeln kan de däremot skilja sig. Elever som har samma betygsvärde kommer också att få samma kalibrerade betygsvärde, vilket innebär att en punkt i praktiken representera fler än en elev.

Figur 1.53 Jämförelse av linjär kalibrering och kvartilkalibrering



Anm.: Blå punkter representerar enskilda elevers betygsvärde respektive linjärt kalibrerat betygsvärde. Gröna punkter representerar enskilda elevers betygsvärde respektive kvartilkalibrerat betygsvärde. Den blå linjen representerar den linjära kalibreringsmodellen. Den röda linjen motsvarar en 1:1 kalibrering /referenslinje).

Vår slutsats är att en stor fördel med kvartilmodellen är att den inte diskriminerar högpresterande elever som går på en skola med en relativt generös betygssättning och där dessa högpresterande elever inte får samma generösa betygssättning som övriga elever på skolan på grund av takeffekter eller att läraren diskriminerar dessa elever systematiskt. Med den linjära modellen syns tydligt i figur 1.53 att de två eller tre mest högpresterande eleverna förlorar på den linjära modellen.

Detta avsnitt ger endast ett exempel på hur det kan slå i ett hypotetiskt fall och fler systematiska analyser, såväl simuleringar som kontrafaktiska analyser baserat på mer omfattande datamaterial måste göras med kvartilmodellen innan man kan säga om den är bättre eller sämre än den linjära modellen.¹ Valet av kalibreringsalgoritm bör inte vara ett politiskt beslut utan bör tas av en expertmyndighet.

Referenser

- Skolverket (2019). *Analys av likvärdig betygssättning mellan elevgrupper och skolor*.
- Skolverket (2020). Simuleringsmodellen och dess kvalitet. Bilaga 3 till rapport 2020:7, *Likvärdiga betyg och meritvärden*.
- Tyrefors Hinnerich, B. & Vlachos, J. (2012). Systematiska skillnader mellan interna och externa bedömningar av nationella prov. Bilaga 4 till *Lika för alla? Omrättning av nationella prov i grundskolan och gymnasieskolan under tre år*. Skolinspektionen.
- Tyrefors Hinnerich, B & Vlachos, J. (2013). Systematiska skillnader mellan interna och externa bedömningar av nationella prov – en uppföljningsrapport. Bilaga 5 till *Olikheterna är för stora: Omrättning av nationella prov i grundskolan och gymnasieskolan, redovisning av regeringsuppdrag*. Skolinspektionen. Dnr U2011/6544/GV.
- Tyrefors Hinnerich, B., & Vlachos, J. (2016). *Skillnader i resultat mellan gymnasieelever i fristående och kommunala skolor*. Rapport nr 2016:10. Institutet för arbetsmarknads- och utbildningspolitisk utvärdering (IFAU).

¹ Det finns även ytterligare varianter på hur kalibreringen kan genomföras som kan behöva undersökas vidare.

Elever och lärares perspektiv på mer likvärdiga betyg och meritvärden

Ungdomsbarometern, på uppdrag av utredningen om likvärdiga betyg och meritvärden (Dir. 2023:95)

Ulrik Hoffman, vd och expert i kompetensförsörjningsfrågor

Jacob Strandell, fil.dr i sociologi och projektledare

Mathias Brodén, mastersstudent och akademisk praktikant

1. Inledning

Utredningen om likvärdiga betyg och meritvärden har i uppdrag att lämna förslag på förändringar i betygssystemet och meritvärdering, syftande till att stävja betygsinflation och öka likvärdigheten i betygssystemet. En viktig del i utredningens arbete är att förstå hur elever och lärare upplever det rådande betygssystemets legitimitet och hur de upplever betyg och bedömning som fenomen, utifrån dimensioner som rättvisa, betygsstress, nationella prov och identitet. Mot bakgrund av detta har utredningen vänt sig till Ungdomsbarometern för att genomföra fokusgrupper på skolor runt om i Sverige.

1.1. Sammanfattning

Sammantaget indikerar undersökningen att det finns flera väletablerade perspektiv och resonemang kring betygssystemet av relevans för utredningens arbete. Eleverna ser instrumentellt på betyg som en valuta för att maximera framtida möjligheter, snarare än som ett steg för att uppnå ett förutbestämt mål. Detta innebär bl.a. att kunskapsinhämtning blir sekundärt och att högre betyg alltid anses bättre. Som mekanism för att öppna upp framtida möjligheter blir maximering av meritvärdet i sin tur det centrala i utbildningsprocessen.

Elever och lärare har olika perspektiv på de nationella proven. Lärarna är generellt sett positiva och ser proven som verktyg för att underlätta betygssättningen, vilken många lärare upplever som svår och tidskrävande. Lärarna är också positiva till att ha fler nationella prov, i fler ämnen, och eventuellt ökad styrning kring hur proven används. Eleverna förhåller sig mer ambivalent till de nationella proven; de lyfter nackdelar som att proven är stickprov vid enstaka tillfällen, upplevelse av stress och ångest inför proven, samt att för mycket vikt vid proven skapar orättvisa för elever som presterar bättre under andra examinationsformer.

Både elever och lärare upplever det nuvarande betygssystemet som orättvist och icke likvärdigt utifrån flera dimensioner. Dessa perspektiv är väletablerade och utbredda, liksom uppfattningen att betygsinflation är ett betydande problem. Trots detta möttes samtliga hypotetiska systemförändringar som föreslogs med varierande grad av skepticism. Detta tycks vara en effekt av att elever och lärare ser betygsinflation som ett samhällsproblem. De oroar sig mer för de konkreta, närliggande eventuella negativa effekterna på individens förutsättningar att få bra betyg respektive farhågan att arbetsbördan för lärarna skulle öka.



1.2. Metod

Datainsamling har skett via sex fokusgrupper fördelade på fyra utvalda gymnasieskolor. Fokusgrupperna genomfördes på respektive skola under skoltid/arbetstid och modererades av en erfaren forskarutbildad projektledare. Vid flertalet grupper närvarade även en utredningssekreterare. Diskussionstiden i respektive grupp var 90 minuter och följde en semistrukturerad intervjuguide som togs fram i samråd mellan Ungdomsbarometern och representanter för utredningen. Samtalen spelades in för att fungera som underlag i analysen, men materialet anonymiserades i analysprocessen för att enskilda deltagare eller skolor inte ska kunna identifieras i slutrapporten.

1.2.1. Urval av gymnasieskolor

Fyra gymnasieskolor identifierades och kontaktades enligt följande urvalsramar:

- Geografisk spridning i Sverige, utifrån en segmentering i fyra segment: Norr, Mellan-Norr, Mellan-Söder och Söder
- Skolor med högskoleförberedande program
- Ej skolor i storstäder
- Ej enbart kommunala skolor

1.2.2. Deltagare

Fokusgruppernas deltagare rekryterades med hjälp av rektorerna på de utvalda gymnasieskolorna, med uppmaning om följande urvalsramar:

- 6–8 deltagare
- Spridning på kön
- Ej enbart elever/lärare från yrkesprogram

Totalt 40 deltagare rekryterades till sex fokusgrupper: fyra elevgrupper (totalt 28 gymnasieelever) och två lärargrupper (totalt 12 lärare). Tabell 1 visar hur deltagarna fördelas på respektive grupp.

Geografiskt segment	Elevgrupper	Lärargrupper
Norr	7 deltagare	
Mellan-norr	5 deltagare	
Mellan-söder	6 deltagare	6 deltagare
Söder	10 deltagare	6 deltagare
Totalt:	28 deltagare	12 deltagare

2. Resultat och Analys

Avsnittet är strukturerat i tre delar som fokuserar på olika aspekter av samtalen: (2.1) *Betyg och bedömning i nuvarande system*, (2.2) *Perspektiv på hypotetiska systemförändringar* och (2.3)



Avslutande perspektiv. De två första delarna summerar tematiskt samtalen i fokusgrupperna. Den sista delen lyfter teman som förekommer genom materialet med de implikationer vi uppfattar som särskilt betydelsefulla för utredningens arbete.

2.1. Betyg och bedömning i nuvarande system

Analysen utgår från intervjuguidens och samtalens första del: upplevelser av betyg och bedömning i det nuvarande systemet.

2.1.1. Betygens roll och funktion

Instrumentellt och med valfrihet som mål

Gymnasieelever ser i stor utsträckning på betyg som en valuta eller ett verktyg för att låsa upp framtida möjligheter, ett perspektiv som framstår som mycket väl etablerat, utbrett och förgivettaget. Betygen ses som ett urvalsinstrument för att maximera framtida möjligheter i livet genom att öppna så många dörrar som möjligt. För att göra det strävar eleverna efter att maximera sina meritvärden. Här är det viktigt att beakta att denna strävan sällan har ett givet framtidsmål, annat än att höga betyg anses ge större valmöjligheter framåt. Det finns således sällan en specifik målsättning, exempelvis att uppnå X poäng för att komma in på utbildning Y. Snarare handlar det om att maximera poängen och därefter välja den mest fördelaktiga utbildningen eller det bästa lärosätet utifrån de alternativ som meritvärdet möjliggör. Detta resulterar i sin tur i att den potentiella stressen kring betygen samtidigt blir oändlig, då det nästan alltid finns utrymme att prestera bättre och få ett ännu högre meritvärde.

Ungdomsbarometern har gott om data från flera tidigare studier om gymnasieelever, såväl kvalitativa som kvantitativa. Dessa studier indikerar att elever generellt sett strävar efter att öppna så många dörrar som möjligt i sina utbildningsval. Detta handlar delvis om att skjuta upp ett val som kan kännas svårt och definitivt. Speciellt på högskoleförberedande program är det mycket vanligt att vilja skjuta upp avgörande val, samtidigt som prestationsviljan är mycket hög.

Det var även förgivettaget bland eleverna i fokusgrupperna att bedömningens funktion är summativ snarare än formativ. Bedömning sågs som mekanismen för att få bra betyg, snarare än som ett verktyg för inläring. I kontrast till det sågs kunskapsinhämtning som sekundärt, som ett medel för att uppnå högre betyg, snarare än som ett mål i sig. Det förekom mycket få resonemang om att samla på sig kunskaper inför högre studier eller arbetslivet. Inga tydliga skillnader framkom mellan programtyperna i denna bemärkelse, men det förtjänar att nämnas att relativt få elever på yrkesprogram ingick i urvalet. Trots detta såg en del elever betyg som en motiverande faktor för lärande, ett slags morot som sporrar dem att lära sig mer, även om syftet med inläringen ändå tycktes vara ackumulering av meritvärde.

Betygens roll för identitet

Eleverna uttryckte sällan att betyg var en del av deras självbild redan vid första frågan, men genom uppföljningsfrågor och implikationer i diskussioner om elevernas förhållningssätt till betyg blev det tydligt att betyg hade viss betydelse för deras identitet. Det framgick då att betyg kan fungera som en identitetsmarkör, och att betygen i viss mån kan användas för att rangordna



individer hierarkiskt, men utan att detta tillskrivs explicita statusaspekter. I stället för generella statusaspekter så betraktades betyg konsekvent som en indikator på personlighetsaspekter kopplade till att vara högpressterande. Exempelvis pratade eleverna om att höga betyg kan betyda att en person är driven, ambitiös, disciplinerad eller smart, medan lägre betyg kan ge omvända associationer, såsom att ses som en "slacker".

Det visade sig dock att vissa elever prioriterade att befinna sig i grupper med andra högpressterande elever av mer pragmatiska skäl. Denna tendens drevs inte nödvändigtvis av en önskan om status, utan snarare av att vilja omges av likasinnade som är lika drivna och kan bidra till att höja den egna prestationen.

Att explicita statusaspekter eller associationer till gruppmedlemskap inte var mer framträdande i grupperna förvånade oss något, då vi vet från tidigare studier att detta beaktas av en del elever vid gymnasieprogramsvalet. Till viss del är detta kanske en effekt av att inga storstäder, där fenomenet är vanligare, ingick i studien.

2.1.2. Upplevelser och erfarenheter av bedömning

Är bedömningens grunder begripliga?

Eleverna upplevde att provresultat ofta var den absolut viktigaste faktorn vid betygssättning, och att enstaka provresultat ibland ensamt kan avgöra betyget.¹ Det fanns även en uppfattning bland eleverna om att aktivt deltagande i klassrummet och i diskussioner kan påverka betygen, men de hade svårt att konkretisera hur detta påverkar bedömningen. Vissa elever upplevde att uppförande kunde påverka ett betyg som ligger på gränsen mellan två nivåer, så att detta justeras uppåt eller nedåt. En del elever upplevde det som orättvist att betyg påverkas av beteende, som aktivt deltagande, snarare än enbart resultat på prov och andra formella uppgifter. Ett exempel som anfördes varför detta kunde vara orättvist var att alla inte är lika bekväma med att tala inför grupp på grund av blyghet eller prestationsångest och därför riskerar att få sämre betyg trots likvärdiga färdigheter.

En annan vanligt förekommande uppfattning bland eleverna var att betygskriterierna upplevdes som otydliga. Eleverna ansåg att lärarna ofta har svårt att motivera sina bedömningar och/eller att ge adekvat återkoppling, samt att lärarna då hänvisar till vaga och "luddiga" kriterier. På en av skolorna berättade flera av eleverna om situationer där lärare inte hade gett någon återkoppling alls eller kopierat ett standardsvar till en stor mängd elever. Denna brist på tydlighet och konkret återkoppling skapade osäkerhet bland eleverna om vad som krävs för att uppnå olika betygsnivåer. Notera här att elevernas efterfrågan på återkoppling framför allt kretsade kring att förstå hur de kan uppnå ett högre betyg, snarare än att förbättra sin kunskap i ämnet.

Lärare: Bedömning är svårt och ökar arbetsbördan

Generellt sett upplevde lärarna bedömning som svårt, med undantag för enstaka lärare med lång arbetslivserfarenhet. Svårigheten var kopplad till osäkerhet i två steg: osäkerhet kring den egna

¹ Här specificerade eleverna inte om det rörde sig om prov generellt eller nationella prov specifikt. Sannolikt beror detta på ämnet i fråga.



bedömningen och osäkerhet kring förmågan att explicit motivera bedömningen vid förfrågan från elever eller vårdnadshavare.

Framför allt upplevde lärare bedömning som svår när det gällde kriterier med värdeord som kräver ett större tolkningsmoment (ofta desamma som eleverna upplever som otydliga, exempelvis är ”nyanserat” ett vanligt återkommande exempel). Dessa kriterier, som involverar ett stort tolkningselement, upplevdes vara svåra att konkretisera om man inte har mångårig erfarenhet av att göra detta.

Många lärare upplevde att det behövs mycket god återkoppling till eleverna för att dessa skulle förstå varför de har fått en viss bedömning och betyg. Kombinationen av osäkerhet i bedömningen och dokumentationskrav gjorde att en del lärare upplevde att de måste lägga mycket tid och arbete på ”bevisföring” i sin bedömning, dvs. att i förväg förbereda motiveringar till bedömningen ifall den skulle bli ifrågasatt.

Lärarna efterfrågade genomgående mer bedömningsstöd för att underlätta själva bedömningen och för att lättare kunna motivera sina bedömningar. Bland annat efterfrågades fler exempel på bedömningar och förtydliganden rörande hur kriterier ska tolkas och användas. Detta skulle i sin tur minska den administrativa bördan, speciellt för lärare under de första fem åren i yrket.

2.1.3. Dagens system upplevs inte som likvärdigt

Eleverna identifierade utan omsvep aspekter av dagens system som gör att bedömningen i praktiken inte upplevs som likvärdig utan skiljer sig åt beroende på omständigheter.

Skillnader mellan skolor

Samtliga elever var överens om uppfattningen att betygsinflation och glädjebetyg är ett stort och utbrett problem, både på grundskole- och gymnasienivå. Eleverna uppfattade att olika skolor sätter olika betyg på likvärdiga prestationer och att vissa skolor sätter högre betyg för att framstå som bättre eller för att locka fler elever. Eleverna kunde dessutom enkelt peka ut skolor där de ansåg att betygsinflation (glädjebetyg) är vanligt förekommande. Kunskap om detta sprids ryktesvägen och via webbsidor för jämförelse av skolor.

Skillnader mellan lärare

Många elever var överens om att lärares bedömningar inte är likvärdiga. Framför allt upplevde eleverna att olika lärare tolkar värdeord i kriterierna olika (här återkommer exemplet ”nyanserat”) och att de därtill bedömer elevers prestationer utifrån vissa aspekter medan andra aspekter missas eller värderas lägre på ett sätt som inte är förutsägbart för eleverna.

Flera elever uppgav att ett centralt element i att få bra betyg är en sorts social konst: att lära sig förstå och spela med i en specifik lärares förväntningar, intressen och värderingar. När man väl har ”knäckt koden” och gör det ”läraren gillar” (exempelvis skriver om ett visst ämne), blir det lättare att få bra betyg. Några av eleverna upplevde även att favorisering av vissa elever förekommer, eller att läraren favoriserar elever som spelar en viss roll (exempelvis ”duktig flicka”).



Skillnader mellan ämnen

Matematik sågs av många som ett mer likvärdigt betygssatt ämne, jämfört med ”mjukare” ämnen som svenska och samhällskunskap. Praktiska och estetiska ämnen i grundskolan upplevdes betygssättas mindre likvärdigt och vara mer utsatta för orättvis bedömning baserat på aspekter som uppförande och favorisering.

Skillnader mellan programtyper

Elever på yrkesprogram upplevde sämre förutsättningar för att få högre betyg i gymnasiegemensamma ämnen. Vid direkt jämförelse mellan elever på samma skola var det tydligt att elever på högskoleförberedande program upplevde sig få mer stöd och återkoppling kring vad som krävdes för att uppfylla betygskriterier inom samma ämne. Elevernas resonerade om att orsaken till detta var att lärare utgår ifrån att elever på yrkesprogram inte vill läsa vidare efter gymnasiet och därför inte värdesätter att få höga betyg i samma utsträckning som elever på högskoleförberedande program.

Lärare: Delade generellt elevernas upplevelser

Lärarna nämnde i varierande grad samtliga punkter som eleverna nämnde, men i mindre detalj och utsträckning jämfört med eleverna. Värt att notera är att vi inte frågade lika ingående om alla aspekter av detta i lärargrupperna, men det tycks som att lärarna i viss mån upplevde samma problem som eleverna.

Framför allt nämnde lärarna skillnader i betygssättning mellan skolor och att fenomen som betygsinflation/glädjebetyg var brett förekommande och allvarliga problem. Vissa lärare pratade även utförligt om skillnader mellan olika lärares betygssättning.

En markant skillnad jämfört med elevernas resonemang var att det stora problemet med betygsinflation i lärares ögon var att det i förlängningen leder till ökad arbetsbörda. Ett vanligt argument var att missvisande betyg från grundskolan resulterade i att eleverna inte har med sig adekvata förkunskaper till gymnasieskolan. Detta kan i sin tur leda till att gymnasielärarna måste lägga extra tid och arbete på att lära ut grunderna och att ge stöd för att alla elever ska kunna ta till sig utbildningen på gymnasienivå. I synnerhet kopplades detta resonemang till elever som gavs betyget E på otillräcklig grund, men i viss mån kopplades det även till andra betygssteg.

Enstaka lärare argumenterade att nuvarande system trots allt är förhållandevis rättvist och betonade att likvärdighet hänger på lärares professionalitet och möjligheter till sambedömning. Då detta inte är självklart i praktiken kan argumentet dock rimligtvis bara tolkas som att likvärdighet upplevdes vara *möjligt* i nuvarande system, snarare än oproblemiskt givet.

2.1.4. Upplevelse av och perspektiv på de nationella proven

Oklar men viktig roll i bedömning

Eleverna hade genomgående i fokusgrupperna uppfattningen att de nationella proven är viktiga för betygen både i grundskolan och gymnasieskolan. Flera elever upplevde de nationella proven i årskurs 9 som stressande och som delvis avgörande för ens slutbetyg i de aktuella ämnena.



Samtidigt var det svårt för eleverna att specificera exakt hur proven användes av lärarna och i hur hög grad proven påverkade betygen jämfört med andra faktorer. Eleverna gav spridda berättelser om att provet hade använts på olika sätt i olika sammanhang. Exempelvis berättade en del elever att provet endast kunde höja betyget eller att provet utgjorde en viss andel av betyget, men dessa berättelser utgjorde aldrig något konsekvent mönster och eleverna verkade ofta osäkra på vad som faktiskt hade gällt. Undantaget var matematiken. Där upplevde eleverna att det nationella provet i princip helt avgjorde betyget. Detta avsåg erfarenheter från årskurs 9, men det fanns även en utbredd uppfattning om att detsamma gällde gymnasiet.

En stressfaktor med risk för orättvisa

Det fanns vissa positiva uppfattningar om de nationella proven, inklusive likvärdighetsaspekter, men över lag pratade eleverna om provens negativa aspekter relaterade till stress och ojämlika förutsättningar att skriva prov. Notera att dessa två punkter går ihop: eleverna upplevde ibland stress på grund av försvårande förutsättningar och den koncentrerade provstressen kring nationella prov kunde i sig upplevas försvårande. Följande tre teman med negativa aspekter av nationella prov förekom i alla grupper:

- **Olika förutsättningar att prestera på prov:** Många av eleverna argumenterade att uppgiften att ”skriva prov” inte passar alla och att elever har olika förutsättningar. Även några av lärarna resonerade på liknande sätt. Som exempel nämndes att vissa elever är mer praktiskt än teoretiskt lagda, att somliga kan ha svårt för att skriva eller koncentrera sig i flera timmar, och att en del är bättre på att skriva inlämningsuppgifter än prov. Några elever pratade om provskrivande i sig som en färdighet som inte alla var lika bra på, inte minst beroende på olika förmåga att hantera nervositet och stress.
- **Ett stickprov är sårbart för ”dåliga dagar”:** Ett annat vanligt tema var att det nationella provet i ett ämne är ett stort stickprov vid ett enstaka tillfälle, vilket gör enskilda elever sårbara för att skriva provet på en ”dålig dag”, dvs. en dag där ens förutsättningar att prestera har påverkats negativt av omständigheter utanför skolan (exempelvis dålig sömn, sjukdom eller stress). En del elever upplevde att detta var orättvist bestraffande på grund av provens stora vikt; en enda dålig dag under hela skoltiden skulle kunna avgöra ens framtid.
- **Resten av terminen, då?** Ett sista tema, som inte var lika frekvent som de andra men ändå vanligt, var att nationella prov kan vara missvisande då de inte tar hänsyn till allt annat en elev har presterat i ämnet, såsom inlämningsuppgifter och muntliga prestationer.

Lärare: De nationella proven är viktiga för betygssättningen

Lärarna i studien var genomgående positiva till de nationella proven och överens om att proven är viktiga verktyg för betygssättningen. Proven upplevdes ge bra och konkreta underlag för bedömning samt något tryggt att förankra bedömningen i, särskilt när bedömningsstöd finns tillgängliga.

Hur proven användes i bedömningen varierade dock. Lärare i matematik uppgav att de använder proven som primärt verktyg för betygssättningen och att de i hög grad sätter betyg som motsvarar resultaten på de nationella proven. Andra lärare uppgav varierande användning av proven, både sett till provens roll i relation till annat bedömningsunderlag och sett till provens



relativa vikt i slutbetyget. En konkret skillnad mellan lärare var att vissa lärare använder proven som utgångspunkt vid sin betygssättning, som sedan justerades av andra bedömningsfaktorer. Andra lärare använde proven som en jämförelsepunkt i slutskedet av bedömningen för att stämna av sin egen samlade bedömning av elevens prestation under själva kursen.

Att proven ska "särskilt beaktas" var därmed något som tolkades på olika sätt av olika lärare. En del av lärarna uttryckte också osäkerhet om hur detta bör göras i praktiken.

2.2. Perspektiv på hypotetiska systemförändringar

Den andra delen av fokusgruppernas intervjuguider fokuserade på möjliga framtida systemförändringar. Dessa förslag på förändringar beskrevs för deltagarna som konkreta hypotetiska scenarion, snarare än som benämnda teoretiska modeller. Rapportens struktur utgår däremot från specifika modellers benämning som var aktuella i utredningen när uppdragets genomfördes.

2.2.1. Fler nationella prov i fler ämnen?

Blandade initiala reaktioner

Elevgrupperna var tudelade i frågan om det bör införas fler nationella prov. Potentiella fördelar som några av eleverna lyfte var: 1) mer likvärdig betygssättning, 2) ökad tydlighet i hur betyg sätts, samt 3) fler möjligheter till konkret återkoppling kring vad som krävs för att uppnå ett högre betyg. Nackdelar som eleverna lyfte motsvarade nackdelarna som eleverna ser med de nationella proven i dag: mer provstress och ojämlika förutsättningar för elever med olika förutsättningar att prestera på avgörande prov (se avsnitt 2.1.4).

Elever vill se fler provtillfällen

Eleverna efterfrågade återkommande fler provtillfällen för de nationella proven för att minska provstressen som annars koncentreras till ett enda tillfälle. Framför allt lyftes två förslag: möjlighet att göra om nationella prov och möjlighet att dela upp proven på flera tillfällen längs vägen i samma ämne, i stället för att ha ett enda stort slutprov. Med fler provtillfällen skulle de generella nackdelarna med nationella prov minska och eleverna var därför mer renodlat positiva till fler nationella prov om det också gavs fler tillfällen.

Lärare: generellt positiva, under rätt förutsättningar

Lärarna var generellt sett positiva till fler nationella prov i och med att de ser de nationella proven som viktiga verktyg vid betygssättningen. Själva tanken på fler nationella prov var därför populär i fokusgrupper, men en del lärare oroade sig för att fler prov också skulle innebära en ökad arbetsbörda. För att motverka det efterfrågade lärarna återigen mer bedömningsstöd, gärna i form av exempel.

2.2.2. Ska nationella prov bedömas centralt?

Eleverna var generellt sett positiva till att bedömning av nationella prov görs av andra än lärarna på den egna skolan, eller centraliseras, för att skapa mer likvärdig och rättvis bedömning. Även



lärarna var generellt sett positivt inställda till att proven bedöms centralt. En invändning från några lärare var dock att de riskerar att tappa det bedömningsstöd som själva bedömningen av det nationella provet utgör. Lärare efterfrågade därför möjlighet att få tillgång till bedömningen av centralt bedömda prov i efterhand för att få värdefull insyn i hur eleven har presterat på specifika frågor och hur svaren har bedömts.

2.2.3. Ska vi ha fler betygssteg?

Blandade reaktioner

Vissa elever lyfte att en fördel med att ha fler steg i betygssystemet vore att dessa lättare kan översättas till precisa meritvärden, men eleverna uttryckte samtidigt också oro för att det kan bli svårt att veta vad som utgör skillnaden mellan olika betygssteg. Eleverna resonerade exempelvis att det skulle bli små nyansskillnader i betygsriterierna som skulle kunna förvärra problemet med otydlighet i bedömningen.

Eleverna hade generellt sett inga starka åsikter i frågan om fler betygssteg under eller över stegen i nuvarande system (dvs. under E eller ovanför A). Några elever såg betygssteg under E som positivt och resonerade att det skulle kunna bli lättare att komplettera kurser eller ämnen från ett betyg som ligger nära E, snarare än att behöva göra om hela kurser. Det förekom även resonemang om att ett betyg som uppvisade prestationer under E skulle ge något att visa upp för eventuella arbetsgivare.

Några elever var även positiva till att införa fler betygssteg ovanför dagens A-nivå, då det kan ge de mest ambitiösa eleverna en ytterligare drivkraft att lära sig mer. Samtidigt tänkte samma elever att det också ökar prestationsrelaterad stress, eftersom man då inte längre skulle vara ”klar” med att ha maximerat betyget med ett A i ett sådant system. Dessa resonemang tycks reflektera elevernas generella vilja att ackumulera maximala meritvärden.

Lärare: Ser praktiska fördelar, men invändningar finns

Flera av lärarna var generellt positiva till ytterligare nyansskillnader inom betygsskalan. Lärarna argumenterade bl.a. att fler betygssteg skulle kunna ge mer fingradig återkoppling till eleverna och att det eventuellt kan motivera eleverna ytterligare. I likhet med eleverna lyfte lärarna att det i dag är problematiskt att elever som får betyget F behöver läsa om hela kursen. Lärarna resonerade att elever som ligger nära E-gränsen skulle kunna komplettera det som saknas utan att behöva göra om hela kursen, vilket skulle vara bra för eleven och kunna reducera lärarnas arbetsbörda markant.

Det fanns dock även flera invändningar mot fler betygssteg. Vissa lärare menade att fler betygssteg än i dag kommer att göra bedömningsprocessen svårare och mer komplex. Detta då det blir ännu fler kategorier att relatera elevernas prestationer till, vilket kan leda till ökad administrativ börda.

En del lärare argumenterade att det i praktiken sällan inträffar att elever inte uppnår endast enstaka kriterier för betyget E, och därmed skulle nära ett E räcka för att komplettera det som saknas i ett system med fler nivåer under E. Det förekom även resonemang om att ”godkänt-gränsen finns av en anledning”; den indikerar att eleverna har tillräckliga kunskaper för att klara av nästa nivå.



2.2.4. Sammanvägningsmodellen

Sammanvägningsmodellen, där lärarbedömningar och nationella prov sammanvägs på ett förutbestämt sätt i slutbetygen, var den modell som elever och lärare hade lättast att förstå. Den uppfattades även ligga närmast dagens praktiker.

Motverkar det betygsinflation?

Några av eleverna var skeptiska till att sammanvägningsmodellen faktiskt skulle motverka betygsinflation. Dessa elever argumenterade att lärare på skolor med hög betygsinflation ändå skulle kunna kompensera för den andel av betyget som de nationella proven avgör genom att sätta ännu högre betyg vid sina egna bedömningar.

Två separata betyg i stället?

I fokusgrupperna diskuterades i viss utsträckning tanken på att ha två separata betyg – ett för nationella prov och ett från lärarbedömningar – i stället för att dessa sammanfogas till ett enda slutbetyg. Elevernas reaktioner på detta var blandade. Elever som gillade tanken såg fördelar i att man i viss mån kan motverka betygsinflationen om det nationella provresultatet tydligt syns, samtidigt som betyget från läraren skulle kunna visa upp förmågor som inte täcks av de nationella proven. Argumenten mot ett sådant system handlade huvudsakligen om att betyget från de nationella proven skulle vara utsatt för de nackdelar som eleverna lyfter med nationella prov generellt, såsom risken för att en ”dålig dag” förstör betyget (se avsnitt 2.1.4).

Lärare: generellt sett positiva

Lärarna var mer entydigt positiva till en sammanvägningsmodell. Lärares positiva inställning berodde främst på förhoppningar om att modellen skulle ge ökad tydlighet i hur proven ska användas, vilket skulle reducera tolkningsmomentet kring vad särskilt beaktande av nationella prov avser. En del lärare såg också fördelar ur ett likvärdighetsperspektiv.

Det fanns dock enstaka invändningar mot tanken på att proven skulle styra betygen på ett förutbestämt sätt som lärarna inte själva kan påverka. Dessa invändningar kom från lärare som ville kunna använda de nationella proven i högre grad än vad styrningen krävde. Den mest utvecklade invändningen kom från en lärare i matematik som i dag använder de nationella proven som en stor del av underlaget för betygssättning och därför oroade sig för att sammanvägningsmodellen skulle kräva att de i stället använder andra bedömningsunderlag i stor omfattning.

2.2.5. Examensprovsmodellen

Reaktionerna på examensprovsmodellen varierade i hög grad beroende på hur ett hypotetiskt scenario formulerades: som ett slutprov i slutet av gymnasieskolan eller som ett antagningsprov för vidare utbildning. Värt att notera är att både elever och lärare hade svårt att föreställa sig och förhålla sig till ett system som skiljer sig så kraftigt från dagens system. Detta antyder att kommunikation, tydlighet och transparens skulle vara mycket viktigt för den här typen av systemförändringar.



Stora slutprov upplevs generellt som negativt

Eleverna upplevde förslaget att ersätta löpande bedömning med stora examinationsprov i slutet av utbildningen som negativt. Ett stort slutprov upplevdes multiplicera de nackdelar som eleverna redan i dag upplever med de nationella proven, såsom provstress, att alla inte har samma förutsättningar för att prestera på prov, att övriga prestationer i kursen ignoreras, och att man blir sårbar för en ”dålig dag”.

Ytterligare en nackdel med slutprov som lyftes av några elever är att elevernas engagemang i sin utbildning sannolikt blir lägre under de första åren, i och med att den skarpa bedömningen skjuts upp till slutet. Dessa elever jämförde detta med liknande erfarenheter från grundskolan, mellan årskurs 7/8 och årskurs 9, och uppfattade det då som att många elever inte ansträngde sig ordentligt förrän det sista året.

Antagningsprov ger mer positiva reaktioner

I jämförelse med slutprov så gillade eleverna i högre grad tanken på att skriva ett antagningsprov efter gymnasieutbildningen för att bli antagen till högre studier, särskilt om dessa antagningsprov utformades för en specifik utbildning eller ett utbildningsområde. Exempelvis föreslog eleverna att man skulle kunna skriva ett prov för att studera till civilingenjör som ser annorlunda ut än ett antagningsprov för att studera till socionom. Detta upplevdes även av eleverna som mer rättvist för att elever med varierande färdigheter skulle kunna prestera bättre på mer relevanta prov.

Lärare: Generellt positiva reaktioner till en roll utan bedömning

Ätminstone hälften av lärarna var positiva till ett system som förskjuter deras uppdrag till att uteslutande handla om undervisning och ”coachning” utan examinerande bedömning av eleverna. Bland de positiva aspekterna med detta nämndes framför allt att det skulle minska lärarnas arbetsbörda. Flera av lärarna resonerade också om att ett sådant system skulle kunna förbättra relationen mellan lärare och elev, liksom mellan lärare och vårdnadshavare, då läraren skulle få en mer stödjande roll inom ramen för undervisningen.

Endast ett fåtal lärare var direkt negativt inställda till en examensprovmodell och kände sig då inte beredda att släppa betygssättningen helt och hållet. Några lärare tog också upp elevernas argument om att slutprov skulle multiplicera problemen som eleverna i dag upplever kring nationella prov.

2.2.6. Direktstyrningsmodellen

Direktstyrningsmodellen knyter de lärarsatta betygen till de nationella provens resultat på gruppnivå (för en klass eller skola). Ett exempel på detta är att kräva att genomsnittet av lärarnas betyg för en grupp elever måste motsvara gruppens genomsnittsprestation på de nationella proven.

Reaktionerna på varianter av direktstyrningsmodellen karaktäriserades framför allt av att modellen var svår att förstå och att den upplevdes vara svår att implementera i praktiken. Modellen kräver därför mycket god kommunikation för att eleverna ska uppleva systemet som legitimt och förutsägbart och omfattande stödsystem för att lärarna ska kunna vikta betygen på ett rättvist sätt utan att det ökar arbetsbelastningen.



Svår att förstå, lätt att missförstå

Moderatorn presenterade direktstyrningsmodellen i form av flera olika hypotetiska scenarion för att illustrera modellen på flera olika sätt, men i samtliga fall krävdes omfattande förtydliganden, exempel och diskussion för att majoriteten av deltagarna skulle förstå modellens generella drag. Trots detta var det vanligt att eleverna, och ibland även lärarna, var förvirrade och missförstod systemets implikationer.

I samtliga grupper associerades modellen till den gamla modellen med relativ betygssättning (med tillhörande missförstånd, som normalfördelning av betygen inom en klass). Flera elever resonerade även om att kravet på att balansera betyg på gruppnivå skulle göra det svårare för lärare att sätta rättvisa betyg och/eller att klassens generella prestationer skulle kunna ha en direkt negativ påverkan på individuella elevers möjligheter att få bra betyg. Detta indikerar att eleverna oror sig för negativa individuella effekter av systemförändringar.

Att modellen var svår att förstå beror sannolikt på att den är den mest komplexa av förslagen och att den ligger långt ifrån nuvarande system. Detta gör i sin tur att modellen kräver mycket kommunikation för att den ska upplevas som legitim och förutsägbar.

Oro för att lärarna inte hanterar fördelningen väl

Även i fallen där eleverna förstod att modellen inte tekniskt sett begränsar individers möjligheter att uppnå ett visst betyg på ett rättvist sätt, resonerade flera elever om att lärarna i praktiken skulle ha svårt att balansera gruppens genomsnittsbetyg på ett rättvist sätt. Eleverna menade att det skulle bli svårt för lärarna att på ett rimligt sätt fördela betygen mellan individuella elever. Vissa skulle då bli lidande. Det förekom också resonemang om att lärare av personlig omtanke skulle vill höja elever från F till E, utan egentligt stöd i prestationen. Då skulle läraren behöva sänka någon annans betyg för att balansera genomsnittet, exempelvis från A till B. Detta upplevdes som mer orättvist än dagens system och som en reell risk, då det ansågs relativt vanligt att höja elever från F till E.

Lärare: Det stora orosmomentet är ökad administrativ börda

Även flertalet lärare upplevde direktstyrningsmodellen som svårt att förstå och drog snabbt paralleller till det relativa betygssystemet och tillhörande missförstånd. En del lärare menade att systemet riskerar att orättvist bestraffa enskilda elever på grund av gruppens prestation.

Utöver att systemet är svårt att förstå var det stora orosmomentet för lärarna risken för ökad administrativ börda i samband med viktningen av genomsnittsbetygen mot ett normerande värde. Många av lärarna oroade sig därför för att systemet vore opraktiskt och komplext att implementera, åtminstone utan avancerade digitala system som hjälper till vid fördelningen av betyg. Vissa lärare oroade sig även för hur fördelningen av betyg mellan lärare skulle gå till och tänkte sig att det skulle leda till förhandlingsituationer som ledde till ökade spänningar på arbetsplatsen och ökad administrativ arbetsbörda för att lösa detta.

2.3. Avslutande perspektiv

I rapportens sista del lyfter vi ett antal perspektiv på anmärkningsvärda och/eller återkommande teman i studiens resultat. Detta är också insikter om elevers och lärares perspektiv som vi på



Ungdomsbarometern, utifrån vår samlade erfarenhet, uppfattar som särskilt viktiga att beakta i utredningens övergripande arbete.

2.3.1. Samhällsproblem ställs mot individers förutsättningar

Lärare och elever var genomgående överens om att systemet i dag är problematiskt och att betygsinflation är ett reellt, utbrett och orättvist fenomen. Problemen upplevdes dock som ett abstrakt samhälls- eller systemproblem. Vinsten av eventuella systemförändringar vägdes i slutändan nästan alltid mot individers förutsättningar. I den ekvationen tycktes samhällsvinster, i termer av ett mer likvärdigt betygssystem, inte kunna uppväga kostnaderna för individerna. Vilka individkostnaderna är skiljde sig åt mellan lärare och elever, men det finns tydliga mönster i deras respektive resonemang:

- Elever oroade sig mer för att individers möjligheter att få goda och rättvisa betyg minskar.
- Lärare oroade sig för att systemförändringar leder till ökad arbetsbörda för individuella lärare.

Dessa teman har två saker gemensamt: 1) motstånd mot förändringar och 2) prioritering av individers förutsättningar framför vinster på systemnivå. Vad detta beror på kan vi bara spekulera om. En möjlig förklaring är att eventuella systemförändringar inte kommer att hinna påverka nuvarande elevers egna möjligheter. En annan är att det finns en trygghet i invanda och inarbetade system som man förstår och har rutiner för att hantera. Men åtminstone i elevernas fall tycks resonemangen motsvara det i hög grad individualiserade och prestationsorienterade perspektiv på betyg och utbildning som vi ser i andra sammanhang (jämför med resonemanget i avsnitt 2.1.1).

Det fanns också en stark aversion mot tanken på att det egna betyget och/eller meritvärdet skulle kunna påverkas av andra elevers prestation, vilket inte minst märktes i diskussionerna om det relativa betygssystemet, eller hur direktstyrningsmodellen skulle kunna flytta en elev från ett rättfärdigt A till ett B, om en lärare samtidigt orättfärdigt höjer en annan elev från F till E. Att betygsinflationen i befintligt betygssystem ger liknande effekter, fast på systemnivå, tycktes inte upplevas som ett lika stort problem.

Oavsett orsak är det här ett perspektiv som bör beaktas. I lärarnas fall kan det exempelvis göras genom att ett nytt system tydligt reducerar arbetsbördan eller att det kommer med så mycket bedömningsstöd som möjligt.

2.3.2. Likvärdighet är inte detsamma som rättvisa

Ett ofta återkommande tema i elevgrupperna var att likvärdighet inte är detsamma som rättvisa. Tvärtom, resonerade eleverna, kan likvärdiga bedömningsomständigheter leda till mer orättvisa situationer för individuella elever. Åtskillnaden mellan likvärdighet och rättvisa var som tydligast i elevernas perspektiv på nackdelarna med nationella prov, baserat på resonemang kring att elever har olika förutsättningar, att stickprov är sårbara för ”dåliga dagar” och att andra viktiga aspekter inte fångas upp av proven. För att ett system ska uppfattas som legitimt är det således viktigt att beakta att det också måste upplevas som rättvis.



2.3.3. Kommunikation kommer att vara viktigt

Ett av de mest konsekventa resultaten var att flera hypotetiska betygssystem upplevs ligga långt ifrån den nuvarande vardagen och att de därför var svåra att förstå sig på och förhålla sig till. Detta ledde ofta till missförstånd och oro för problem; problem som inte nödvändigtvis var effekter av den hypotetiska förändringen. För att undvika missförstånd kommer därför kommunikation om systemförändringar att vara mycket viktigt för systemets legitimitet hos både lärare och elever. Minst fyra aspekter måste beaktas och kommuniceras kraftfullt och i stor omfattning till både lärare och elever om större förändringar görs:

- Hur fungerar systemet (försäglighet)?
- Varför görs förändringen/Varför är det viktigt?
- Vad gör det för skillnad i praktiken?
- Hur ska det nya systemet användas?

2.3.4. Skillnader mellan program

Även om betyg är en fråga som engagerar de allra flesta eleverna så finns stora skillnader mellan grupper i fråga om hur viktigt de upplever att det är att få bra betyg, deras kunskaper om betygssystemet och deras förmåga att sätta sig in i hypotetiska förändringar. En skillnad i studien mellan högskoleförberedande och yrkesprogram var att eleverna på yrkesprogram ibland förväntades vara mindre intresserade av betyg. Samma tendens finner vi också i våra andra studier. Föreställningen stämmer dock inte alltid överens med verkligheten och riskerar då att leda till lägre förväntningar på och mindre stöd till dessa elever.



CAMBRIDGE

Improving the Comparability and Reliability of the Sweden National Tests

January 2024

Prepared by Matthew Dean [Cambridge International Education]
Georgie Billings [Cambridge International Education]

Date January 2024

Transforming societies
through education

Table of Contents

Introduction	2
Purpose and content of the report	2
About the authors	3
Summary of context	3
Control factors	3
Explanatory factors	6
Factors that can affect test reliability	7
Analysis and recommendations	7
Assumptions	7
1. Test construction	8
2. Checking the target standard	9
Expert panels	10
Subject Pairs analysis	10
3. Carrying a standard between test versions	10
Norm referencing (fixed cohort distributions)	10
Fixed thresholds	11
Qualitative standard setting	12
Quantitative methods	13
4. Test equating	16
Classical Test Theory	16
Rasch / IRT models	17
Overall	18
5. Scaling and reporting	19
Role of the national test scores in students' final outcomes	19
Implications for security and administration	20
National tests for adult education	20
Bibliography	21

Introduction

This report has been produced by Cambridge University Press and Assessment in collaboration with the Swedish Government Commission into *Maintaining Equivalent Grades and Credits in Secondary Schools* (*Utredningen om likvärdiga betyg och meritvärden*). The report will be published as part of a wider government-commissioned report on the same theme in 2025.

Purpose and content of the report

The agreed purpose of this report was to make recommendations that will improve the consistency and construct coverage for the Sweden National Tests for the lower secondary (Grade 9) and upper secondary levels, with a view to the results of these tests then being given more weighting in the overall grading of students.¹ Identified parameters for this piece of work included:

- the standards set and carried must be criterion-linked in some way, either qualitatively or quantitatively.²
- the suggested methods should broadly work for the required assessment levels and subjects, though they may differ for different levels as appropriate.
- the methodology should have face validity and be communicable to stakeholders. This will involve linking it to research, usage and literature in the field of assessment.

- the methodology should allow for results to be comparable between administrations (including tests from different years) and between subjects.
- the methodology should take into account the practicalities and resources involved, such as cost, scale, expertise and time.

A joint piece of work over four months was planned to meet the following objectives:

1. Explore possible options for standard setting and maintenance, including challenges and a benefits analysis of each.
2. Deeper examination of preferred options for standard setting to establish technical requirements.
3. Production of a report which makes a reasoned recommendation for moving forward.

The work started in November 2023 with three experts from Cambridge spending one week in Stockholm to gather information and data from relevant stakeholders including the members of the commission. Cambridge then compiled a report based on the findings of the initial data gathering stage which was submitted to the commission in December 2023 for review and feedback. The final report was delivered in January 2024.

¹ Some of the recommendations from this report may also be relevant for the compulsory tests that take place in Grade 6.

² The term 'standard' as used in this report primarily refers to the *level of demand* of the National Tests as opposed to, for example, a reference to the *content standard* of the tests (i.e.

the curriculum content). The terms 'standard setting' and 'maintenance of (or maintaining) standards' refer to the setting of *performance standards* of students in the different test administrations. For a fuller discussion of these terms and how they are used, see Baird et al (2018, pp. 54-57).

Cambridge previously produced a report relevant to this work in 2018 in collaboration with the Swedish National Education Agency (*Skolverket*) entitled *Marking and grading models in Swedish National Tests — An analysis and recommendations for future models*.³

For this report, reference has been made in places to arrangements in the UK, and more specifically to England's system of curriculum and assessment (secondary level). This is because the authors feel there is a useful level of convergence and divergence between the two systems. For example, high-stakes national tests take place in both systems at roughly the same point and the tests themselves are comparable in design and quality. At the same time, there are some fundamental differences in curriculum content and test development, production, administration and reporting processes.

About the authors

The authors of the report are part of Cambridge International Education's Assessment and Customer Support Directorate, and work in the Assessment Projects Group.⁴ This group has been working on education reform projects with a wide range of partners from around the world since 2011, including education ministries, NGOs & donor organisations and private school groups. The group works closely with Cambridge Partnership for Education to deliver educational assessment-related projects.⁵ All of the

above are part of the University of Cambridge.

Summary of context

In the area of education reform, a *control factors* analysis has been developed by Cambridge over the past decade which builds upon prior research on 'curriculum coherence', 'curriculum control', an ongoing review of transnational research, and practical development work with education ministries and school systems in a range of nations.⁶

Approaches should take into account the specific pressures and possibilities which make up the context in each national setting. Many variables are at play within the process, all of which, singly and in interaction, have some effect on educational system effectiveness, cost effectiveness and ultimate quality. Some influencing variables are within government control and can in principle be modified by policymakers as they consider appropriate, to improve their education system through anticipated impacts in the short, medium or long term.

Control factors

For this report, we have used these control factors to broadly categorise and present some of the most relevant and contextual aspects of the current educational context to this proposed reform measure⁷:

³ This report made recommendations related to the rationalisation of the number of grading models across subjects in operation at the time.

⁴ For more information on Cambridge International Education, see:

<https://www.cambridgeinternational.org/>

⁵ For more information with Cambridge

Partnership for Education, see:

<https://www.cambridge.org/partnership>

⁶ Schmidt, W., & Prawat, R. (2006). Curriculum coherence and national control of education: issue or non-issue? *Journal of Curriculum Studies*, 38(6), 641–658.

⁷ See the body of the government report for a fuller description of the context to this proposed reform.

Curriculum content

The last major reform to the Swedish National Curriculum took place in 2011. In terms of subject specifications and (subject) syllabus content, it describes what (knowledge) teachers should teach and students should learn at a relatively high level ('core content') when compared to, for example, the National Curriculum for England.⁸

Of particular importance for the purposes of this report are the subject-specific (content) standards and grade descriptors (*betygskriterier*) which can be found in the Swedish National Curriculum document⁹, that are used by teachers when awarding students a grade at the end of a school year or upper-secondary course. Currently, the national test result for a student needs to be considered but the teacher is ostensibly free to award a grade different to the test outcome.

These statements describe what a student at a particular level of ability (key grades at 'A', 'C' and 'E') should know and be able to do. Award of a Grade 'F' in one or more subjects is of particular significance as it may lead to remedial action and/or limit a student's progression into upper secondary programmes.

Changes to make the (content) standards more specific and less subjective have been made in recent years in an attempt to improve the reliability of student grades between schools and test administrations (conducted at different points of time) but these changes have not had a notable impact.

⁸ See [The national curriculum in England Key stages 3 and 4 framework](#). Individual examination boards produce subject specific syllabuses based on the National Curriculum which breaks down the subject content into specific *learning objectives* which can then be used to map examination

Pedagogy

Teachers are involved at various points during the test development and production process. Individual university departments are responsible for managing teams of teachers who take part in standard setting exercises, the writing of items/tasks and advise on the coverage and representation of curriculum content.

The National Tests are seen as helpful to the teachers involved as the test content can affirm and have a positive *washback effect* on their classroom practices.

Assessment and qualifications

Currently, Swedish primary and secondary school students are academically tested at various points of their school careers, in a mixture of compulsory, whole-cohort tests, (where all students are tested in a subject), and sampled subjects (where a representative group of students are tested). There is a low level of consistency in test scores and grades between test versions.

National Tests in scope for this report include those taken at the end of Grade 9 (the end of mandatory education) and those currently being administered during the upper secondary stage. Currently only Swedish/Swedish as Second Language, mathematics and English are tested in this way and these tests only account for 10-15% of the upper secondary certificate. Note that vocational assessments and qualifications are out of the scope of this report.

content between test versions and years (learning objectives are listed in *setting* and *coverage grids* during the test production process).

⁹ [Curriculum for the compulsory school, preschool class and school-age educare](#)

Each national test is produced by a different university department, marked by teachers, and then each student is given a test grade in each subject. The standard setting for the test currently takes place before the tests are taken.

The Swedish National Education Agency is responsible for overseeing the development and production of the National Tests and produces key reference documents including a framework document for the production of subject tests.

While some efforts have been made to standardise the development and quality of the tests, there is limited coherence in how standards are set and carried.

The purpose of the tests is to inform the final grade allocated to a student, but this is not formalised and teachers are free to award the final grade for each student based on their own estimation of the individual student's level of knowledge and ability as rated against the grading criteria for that subject.

Currently teachers mark their own students' assessments, but there are plans to centralise marking with a standardised pool of teachers, following the move to on-screen assessment and digital marking. This move will also facilitate more statistical analysis that may have a positive effect on the quality of the tests and reliability of outcomes.

Question papers are not regularly released though sample items for individual subjects are available. Items are regularly withheld for possible re-use for up to six years.

Governance & Accountability

After the establishment of free schools (*friskolor*), which are state-funded privately-run schools, there has been

some speculation that grades for students are not reliable and a perception of *grade inflation* across the school system as a whole.

Other factors that may lead to this perceived grade inflation include:

- the desire of some teachers to help certain students
- teachers with limited experience of teaching and/or of assessing a wide range of students in terms of ability, inadvertently advantaging or disadvantaging individual students when awarding final grades.

Professional development

Teachers directly involved in the test development and production process benefit in terms of assessment knowledge and practice (see *Pedagogy* sub-section above).

Inspection

National Tests are administered by schools and the schools have a level of autonomy in how the individual subject tests are administered. Test papers are delivered by post/courier service and there is a test administration window. There is no dedicated system of school inspection in place around the test administration in schools to help ensure that schools are following regulations and best practice.

Funding

Changes to the production of National Tests are likely to incur new costs that will need to be met, e.g. commitments to onscreen assessment and centralised marking (already committed). Further changes to the National Tests, including for example more post-hoc centralised processes requiring specialist input and human resources, may depend on the availability of funding.

National framework

At upper secondary level, there are a number of *programme* options leading to an upper-secondary certificate for an individual student.

The number of national tests a student will take depends on which programme (e.g. *vocational, natural science* etc.) they take, and the individual *learner journey* to the certificate in terms of required courses and assessments is relatively complex. Performance on each course earns an individual student a number of merit points leading to the awarding of the certificate and facilitating entry to higher education. As is the case in Grade 9, teachers take into account an individual student's national test result when they award the final grade.¹⁰

Selection and gatekeeping

From their third year at upper secondary level, students can also gain entry to higher education by taking an aptitude test that is held twice a year and assesses verbal and mathematical reasoning. The SweSAT Scholastic Aptitude Test (*Hogskoleprovet*) is administered by the Swedish Council for Higher Education¹¹.

Explanatory factors¹²

Some of the variables that affect the structure of an education system and its functioning are beyond the control of governments. Such 'explanatory factors' of particular relevance to this report include:

¹⁰ As a consequence of this structure, students are not always assessed at the end point of secondary education for some of the subjects that are required for admission to higher education courses.

¹¹ [https://www.uhr.se/en/start/system-support-services/system-support-services-for-higher-](https://www.uhr.se/en/start/system-support-services/system-support-services-for-higher-education-institutions/registering-for-the-swedish-scholastic-aptitude-test/)

Culture

The fact that for many years students' teachers have been responsible for awarding their own students' final grades has had a role in shaping current *societal attitudes* in this regard.

Reforms to the education system have seen a move towards criterion-referenced assessment whereby students are measured and outcomes reported against set criteria.

However, recent performance by Swedish students in international tests has not always matched the outcomes of the National Tests, even when demographic changes and events such as the Covid-19 pandemic have been accounted for.

Political structures and commitments

Different approaches to how the national tests should be used have been adopted by different government administrations. In recent years, stated multiple purposes of the tests have been refined to a single main purpose, namely to support teachers in the grading of their students.

Historical contingencies

Large-scale demographic changes that have taken place in recent years are likely to have had an effect on the test results when viewed at a cohort level. One such change includes an increase in the level of inward migration and the time required for new arrivals to adapt to a new learning environment as well as to become sufficiently proficient in the Swedish language to enable a valid and

[education-institutions/registering-for-the-swedish-scholastic-aptitude-test/](https://www.uhr.se/en/start/system-support-services/system-support-services-for-higher-education-institutions/registering-for-the-swedish-scholastic-aptitude-test/)

¹² For the full list of *control* and *explanatory factors* and a description of each, see [Cambridge Approach to Improving Education](#). 2017, p.20.

fair assessment of their knowledge and ability.

Factors that can affect test reliability

Some relevant and more direct factors that may have an effect the reliability of the outcomes of the national tests (i.e. cause *construct irrelevant variance*), and which are not covered directly in the scope of this report, include:

- **administration** - as there is a window of time for individual schools to administer each assessment, there is the possibility that some students may be able to receive information on the test content in advance.
- **inspection** - The lack of school inspections during the examination period means that there is limited information and data on school compliance in how question papers are stored, examination sessions run etc.
- **marking** – students' test papers are routinely marked by their own subject teacher and most responses are not anonymised. Even where they are anonymised, such as for the essays in language examinations, teachers are frequently able to identify students based on their composition style. It is relatively unusual for this approach to be used for such high-profile and high-stakes assessments. The planned move to centralised marking will facilitate anonymous marking thus removing perception of associated bias.

Analysis and recommendations¹³

Our analysis and recommendations cover five areas all directly related to setting and maintaining test standards (or alternative methods) across multiple test administrations and test versions.

These are presented in a sequential order in terms of the test development, production, administration and reporting process:

1. Test construction
2. Checking the target standard
3. Carrying a standard between test versions
4. Test equating
5. Scaling and reporting

Assumptions

- It is assumed that the general approach to testing will be reviewed prior to implementation of any recommendations. This may involve rationalisation of tests, introducing testing in more subjects with more frequency and/or re-setting of the test specification and standard.
- It should be noted that the option to conduct sample-based assessment (where only selected students take the test, or where the test is not offered each semester) to inform overall grades is not appropriate. This is because students who sat the test, and therefore had to inform their grades, would be treated differently to those who did not. Some further modelling to predict the correlation between test scores in other subjects

¹³ This section assumes some knowledge of assessment-related methods related to test *equating* and *standard setting*. For a helpful paper that helps to explain these methods and also compares the methods on two tests, see:

<https://www.cambridgeassessment.org.uk/Images/424273-comparing-small-sample-equating-with-angoff-judgment-for-linking-cut-scores-on-two-tests.pdf>.

and achieved grades in others would be needed to design any such model.

1. Test construction

It is desirable to have test versions that are designed to be as equal between sessions in terms of difficulty as possible. This provides a more consistent experience for test-takers (important for reducing test-irrelevant performance factors) and minimises the movement of raw score thresholds.¹⁴

At present, Angoff exercises are carried out prior to the live test administration to gauge comparative test difficulties. However, the fluctuating results of the scores in Maths and Sciences (see Figure 1) indicate that this is not always successful in producing equivalent tests (there is no particular reason why a single-country cohort would perform this differently year-to-year).

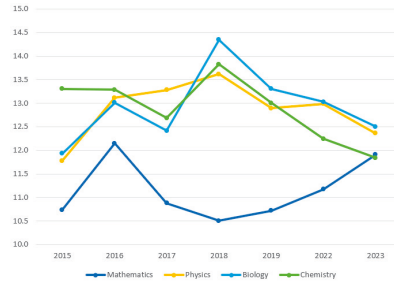


Figure 1: Test grade points for maths and sciences (2015-2023)

Recommendation 1:

Cambridge recommends that a more detailed test specification is designed by the test constructors and used as a guide for each test.

There is no need for this to be uniform across subjects and levels, but once established for a particular subject and level, it should be used consistently for every test version across a number of years.¹⁵

Suggested areas for inclusion in a test specification are:

- total number of raw marks to be included in each test
- question types

¹⁴ Currently some versions for the same subject test have a different number of total marks.

¹⁵ Examination Boards in the UK, under the guidance of the Federation of Awarding Bodies

(FAB), routinely review test specification documents about every 4-5 years.

- any weightings within or between tests
- the broad brackets for skill taxonomy coverage (e.g., 10-15% of marks assess knowledge, 20-30% of marks assess analysis etc.)
- the broad brackets for curriculum content coverage (e.g., 20-25% algebra, 15-20% trigonometry etc.)
- the broad brackets for estimated item difficulty. These can be expressed in a variety of ways, such as:
 - facility value brackets
 - target grade ability or ‘easy, medium, difficult’.

The rationale for this is that inconsistencies in the design of the test, including the total number of marks, is likely to cause swings in the distribution of candidate scores (not simply the overall difficulty of the test) and may necessitate re-establishment of the standard from scratch each year. Such inconsistencies also increase the likelihood of test-irrelevant factors impacting on candidate performance (fewer/more marks than expected, different item types/content than expected etc.).¹⁶

The key to establishing a stable test difficulty and therefore standard over time is to ensure consistency in setting. Failure to do this will make using anchor items fundamentally unviable, as the distribution, quantity, and characteristics of these would also therefore be unstable, as they must be reflective of the overall test.

¹⁶ Another key resource that helps to ensure test uniformity is a document with detailed instructions for the writing and production of content for a specific test which is made available to setters and item writers (often referred to as *Item Writer Guidelines*). This document can also include

Of equal importance is that these specifications, once designed, are vigorously adhered to by test developers, who should map each test to the specification. A record of these can provide valuable insight into test coverage (topic, item type, skill, difficulty level) over time, especially when compared to item-level data from live administration.

When the tests are administered online, item-level performance data on the items should be studied by test developers against their estimated difficulties to understand how items perform and can be improved.

2. Checking the target standard

Recommendation 2:
Before implementing any new standards maintenance measure, Cambridge recommends that a thorough review of the existing standards across all tests is undertaken to ensure that the standard from the previous administration is actually the standard that should be maintained.

There are a number of possible ways of implementing the recommendation

examples of well written items and tasks as well as examples that have been shown not to function as intended. This document should be continually updated and refreshed and reflect any changes to the curriculum.

above, but Cambridge offer the following for consideration:

Expert panels

This involves gathering subject teachers and sharing with them some anonymised examination papers of test-takers in the previous administration who scored just enough marks to achieve a grade of A, C and E. These should be compared to the written standards. If the judges feel these do not match, papers at higher and lower marks can be shared until a mark is found that matches the written standards. This can then be seen as the ideal standard to be carried forward to the next test administration or over a longer period of time.

Subject Pairs analysis

Any different subject tests that are seen as potentially carrying an equal number of credits should be compared, on the premise that it should not be easier to get an 'A' in one subject than another (while individuals will of course find certain subjects easier than others, across the cohort as a whole this would be expected to smooth out).

To do this, a sub-set of test-takers is identified who sat a pair of subjects (for example, all the students who sat maths and chemistry exams at the same level).

By comparing the scores of just these paired test-takers across the two subjects, we can see whether they are broadly aligned (for this purpose, a variance of 5% or less in grade outcomes can be considered well-aligned).

3. Carrying a standard between test versions

The objective is to carry the performance standard from the previous test version whenever a new test version is used. This essentially means that students who perform equally well should receive the same grade, no matter when they take the test.

Norm referencing (fixed cohort distributions)

We do not discuss here the various ways of cohort (or 'norm') referencing at length (also referred to as 'fixed distributions'). This method was used in Sweden previously but would not necessarily help achieve the stated objectives of the tests as they stand now (that is, to support the grading of candidates).

Under this method, the distribution of candidates would be kept stable every year (so, for example, the mean mark and standard deviation could be used to force the distribution into the same shape each year or be used to move cut scores around to keep the proportions of candidates at certain grades stable).

The key weakness of this approach is that one cannot be certain that the overall cohort of candidates is identical in distribution to the previous year, which is difficult to prove empirically. Similarly, it is unsuitable for tests that are administered on-demand since the candidature would be different in each administration.

In reality, this approach would mean that a candidate's grade was decided by how good other test-takers in that administration were compared to that person: if the population was generally better, then it would get harder for that person to get into the top bracket of A-

grades, since only a limited proportion of test-takers will get through.

If the only objective of a test is to rank order the candidates who took it in that single administration of the test, then norm referencing profiles a readily understandable and simple to administer mechanism. However, there are some drawbacks where you wish to be able to compare between administrations of the test.

For example, a student who takes a test one year may get into the top 10% of candidates, and then defer their place at a university for one or more years. The following year, the cohort may be better overall, and so the performance required to get into the top 10% of candidates would be greater. This would mean the student who deferred is advantaged, as they did not have to perform as well as students from the following year.

Another example might be if you wish to use the test to measure improvements or declines in standards over time: as norm referencing is explicitly about keeping the population distribution equal year-to-year, it offers no insights here. With the public perception of declining standards, this could be seen as a disadvantage in the context of Sweden.¹⁷

Fixed thresholds

Where there is high confidence that specific subject tests or test components are of equivalent difficulty between test versions, the same thresholds or scale can be used.

There are some instances where this is appropriate. Examples might be essay or long-form examinations, where the test questions are used more as a prompt to elicit a demonstration of skills from students, and the rubric (mark schemes) and marking standard remain the same. Examples of this may be speaking or writing examinations in languages.

Another example might be tests that have been constructed from pre-calibrated items (e.g. items from pre-testing activities or from previous test administrations) to ensure they are of exactly equivalent difficulty to the previous version.

Recommendation 3: Cambridge recommends that the fixed threshold approach continues to be used for English, Swedish and Swedish as a Second Language, where the data supports that the demand standard of the test has been steady over time (see Figure 2 below). If the format of these tests is revised, then this decision should be revisited.

¹⁷ A form of fixed distribution (commonly referred to as 'comparable outcomes') is used in the UK for some qualifications, including the high-stakes GCSEs (national examinations taken by students at the end of Year 11). For a description and

further discussion of the benefits and drawbacks of this approach, see: <https://www.cambridgeassessment.org.uk/Images/346267-comparable-outcomes-scourge-or-scapegoat-.pdf>

Improving the Comparability and Reliability of the Sweden National Tests 12

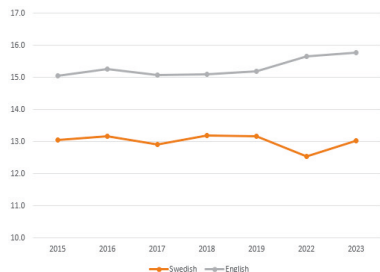


Figure 2: Test grade points for languages (2015-2023)

Qualitative standard setting

At present, the standard (cut scores for each letter grade) is set using an Angoff methodology before the live administration of a test. This involves expert judges estimating how each item in the test will perform at each grade. For example, a judge might be asked how many individuals in a group of minimally competent A-grade test takers would get a question right. These are then averaged for the different judges, or discussed until a consensus is reached.

While this can be a robust method for maintaining standards, it is clear from the performance statistics for some of the current tests that this is not always working, with large swings in outcomes between years which cannot be explained by differences in the cohort distribution indicating that the issue is in the varying difficulties of the tests (see Figure 1 above and Figure 3).

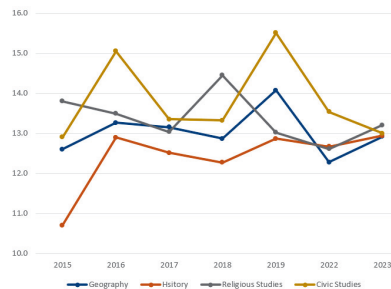


Figure 3: Test grade points for social science subjects 2015-2023

Relevant factors when using the Angoff method can be seen in Table 1 below:

Challenges	Benefits
Involves pre-exposing the test content to a small group of experts: possible security concern	Involves the subject community in the maintenance of standards, helping face validity
Logistically burdensome to recruit and standardise sufficient judges, arrange meetings, gather and process data etc.	A common qualitative method in the literature
Expensive and time consuming: several experts are required over several days for each subject and level	The standard can be set before live administration of the test
Mixed results reported in terms of accuracy of this method	Statistically easy to calculate

Table 1: Challenges and benefits of using the Angoff method

Recommendation 4i:

If Angoff continues to be used as the primary method, Cambridge would advise that a robust set of guidelines is produced to ensure this exercise is carried out consistently and to the highest standards.

These guidelines should include:

- instructions about the composition of the judging panel (number of judges, experience required, exclusion of any judges with a vested interest in the test, such as setters)
- instructions about standardising (training and co-ordination) the judges. This will involve ensuring the judges have a clear understanding about the capabilities of a minimally competent test-taker at each grade. This can usually be gathered by using a previous test along with test-taker data. For each item, a facility is calculated from just the test-takers who were only one mark over the threshold for each grade. A single item will therefore have five facilities, descending in value (e.g., A facility 0.78, B facility 0.64, C facility 0.49, D facility 0.36, E facility 0.28). Then, for each grade, those items with a facility of around 0.5 are identified as being 'boundary items': items that are at the edge of the ability of a minimally competent test-taker (in the example above, the C facility of 0.49 means it would be tagged as a boundary item for that grade). Taking each grade in

turn, these items would be discussed amongst the judges before they are asked to make their estimations for items in the new test at that same grade.

- instructions for the estimation method (e.g., *yes/no*, *facility prediction*, *modified Angoff* etc.). At the moment a variety of methods are in use. In at least one case, it seems that a variety of methods are used for a single test. This will not lead to increased precision, and indeed the evidence that different methods give different outcomes indicates that Angoff is not working well at present.
- instructions for aggregating judges scores. Where judges do not agree, should individual judges' estimations be averaged? Should outliers be accounted for in some way?

Recommendation 4ii:

Cambridge recommends that Angoff is not used as the primary standard-setting method, due to concerns about accuracy and operational viability.

It could instead be used periodically as a check on a standard set by statistical equating. This would alleviate concerns about any perceived standards drift from use of a statistical method.

Quantitative methods

The assumption is made that the same test-takers are unlikely to take multiple versions of the same test in any significant numbers, and so common-cohort equating is unlikely to be possible. This leaves us with common-item equating.

Pre-testing whole test

A statistical pre-test can be carried out to pre-calibrate all the items in a test (we understand that a large amount of pre-testing is currently carried out by the universities who produce the tests but that items are not calibrated as a result of these pre-tests).

Note this is distinct from an item validation trial of items, which would typically involve lower numbers of test-takers and have the objectives of just ensuring that the item (or item type) is broadly functional. Also note that whole tests do not need to be pre-tested as a single test; items can be pre-tested in different batches.

Such large-scale statistical pre-testing is challenging in many ways. See Table 2 below for a summary of the challenges and benefits:

Challenges	Benefits
Test-takers are often not motivated to try in these pre-tests.	The greatest advantage of this is that it allows the standard to be set before the test is administered.
Involves pre-exposing the test to test-takers	
Can be difficult to recruit a demographically representative sample each time	
Must be taken by test-takers at the same academic point in their journey. This usually means sitting it alongside their live test, leading to an increased burden on students.	
Can be difficult to recruit test-takers in sufficient numbers	

Table 2: Challenges and benefits of pre-testing whole tests

Anchoring

Anchoring involves the inclusion of a small proportion of pre-calibrated items. Performance in these items is then used to calibrate the rest of the test, which may include items that are new.

This has a number of advantages over pre-testing the whole test, including a reduced impact on test-takers, better quality data, and lessened security

concerns. It should be noted that unless all items in a test have been pre-tested or pre-exposed, it will not be possible to set the standard in advance of the live administration: they will instead be used as anchors in a post-hoc statistical equating of the live test. Because the psychometric properties of anchor items must be known, they are usually items that have been pre-tested or used in a previous test (a 'pre-exposed' item).

Anchor equating can be broadly classified into two types:

1. The first approach to using the anchor scores is *chained equating*. It consists of equating the scores on the new form to scores on the anchor and then equating the scores on the anchor to scores on the reference form. The "chain" formed by these two equatings link the scores on the new form to scores on the reference form.
2. The second approach involves conditioning on the anchor/post-stratification. Here, the anchor score is used as a predictor variable. For each possible score on the anchor, the distribution (or possibly just the mean and standard deviation) is estimated of scores on the new form and on the reference form in the target population.

Pre-tested items:

Some new items are included in a live test which will not count towards a test-taker's final score or grade (*external*): they are items that will be used in future test versions. This allows them to be pre-calibrated.

Pre-exposed items:

A proportion of items in a test will be pre-calibrated by prior use in a live test. These items should represent a balance of curriculum coverage, difficulty and item type.

They can be used as part of the test-takers' final scores/grade (*internal*) or used only for calibration of the test (*external*).

The anchors can be drawn from any item for which a calibrated item difficulty and discrimination is known, and which functioned well. This means that over time, you will build up a large bank of potential anchor items which can be drawn from multiple exposed tests, reducing security concerns.

Recommendation 5:

Anchor equating should be used as a more general principle that can easily incorporate many of the method-specific advantages discussed here.

A minimum of 10% or 10 test items should be anchors and these should, as far as possible, include a range of curriculum content, item types and item difficulties.

However, it may be desirable to have more items at areas of the distribution that are of particular interest (for example, if the boundary between 'F' and 'E' is where you need the most precision, it may be wise to focus more anchor items on this type of test-taker).

The tests could be extended to allow for these, or they could be incorporated within the current test length. These can be internal or external to test-takers' scores/grades. Ideally, they should be distributed through the test as they would be in any other test and tagged as clearly identifiable in any data streams.¹⁸

4. Test equating

Assuming that at least 10% of items in a test are pre-calibrated, anchor items (and there is a minimum of ten such items), then the test versions can be statistically equated to ensure that any variation in test difficulty is accounted for.

There are many ways of doing this, but here we outline some possibilities. In general, the choice will be informed by the competing considerations of absolute accuracy and capacity. It may be that the most robust equating model is not the most appropriate, as there is not sufficient psychometric resource available to maintain accuracy levels.¹⁹

Classical Test Theory

The approaches below are based in Classical Test Theory (CTT) statistics, such as facility and discrimination. In equating, the tests are viewed as a whole, rather than at item level: while item level statistics can be calculated, it is the aggregated whole which is then used in equating. CTT has the advantage of being easy to compute and use.

Mean Equating

Principally, this equating method compares the confidence interval of the test mean with the equivalence limits of the reference mean. Two scores

are deemed equivalent if one score's test mean falls within the equivalence limits of the other. The benefits and limitations of this method are presented below:

Limitations	Benefits
In a classroom, it may be that other descriptive statistics are just as important.	Ease of explanation
The impact of any non-standard classroom cohorts can quickly 'become lost in the data'. For example, the impact of any semester's under- or over-performing cohorts would be made much less visible by using mean scores – essentially information at the tail of the distribution is less accurate. Where 'E' is seen as such a crucial grade, this may be a significant concern.	Minimal resourcing requirements
	Can be computed easily

Table 3: Limitations and benefits of mean equating

Equipercetile Chain Equating

Here, we would look at the mean score of a test-taker across the anchor items in each percentile in the new test. Any level of granularity could be considered (e.g. 25%, 50%, 75%, or 10%, 20%...90%, or 1%, 2%...99%). We would compare this to the mean anchor scores of test-takers

Test Equating, Scaling, and Linking: Methods and Practices (Statistics for Social and Behavioral Sciences), Springer

¹⁸ This refers to adding key attributes to each item according to a pre-defined taxonomy.

¹⁹ For more information about methods for equating, see: Kolen M.J & Brennan L. (2014).

in each percentile in the reference forms of the test.

For example, if in the reference forms of the test, test-takers in the 90th percentile scored on average 1.2/5 marks on the common items, but in the new form they scored on average 1.4/5 marks on the common items, we can infer that the test-takers in the live test are slightly better than the test-takers in the reference versions: the items are of the same demand, so improvement is down to the quality of the cohort.

We can then look at the overall scores for test-takers in that percentile for the new and the reference forms of the test. If overall scores are similar, then the test must be more difficult for the test-takers in that percentile, as we know they are more able, and so would do better in a test or equivalent difficulty.

By plotting this for test-takers in each percentile, we can calibrate the scores of the new test form to match the reference test form. This example is represented graphically in Figure 4.

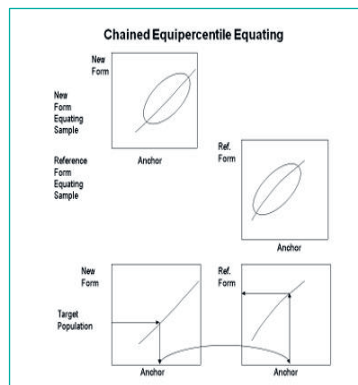


Figure 4: Illustration of chained Equipercentile Equating²⁰

While more complex, this can still be calculated relatively easily without specialist software or psychometricians and is easier to present in graphic form to stakeholders.²¹ It also allows for the test to vary more or less in difficulty depending on the overall ability of the test-taker (for example, a weak test-taker may find the new test harder than the reference form, but able test-takers may not), allowing retention of distribution definition at the tails (a major issue in mean equating).

Rasch / IRT models

This section will not attempt to explain all the possible variations within these models or the underlying algorithms; this is a very technical area and is better explained in the sources listed in the bibliography for those who wish to engage in more detail (see page 21). Instead, it gives an overview of the principles, and the potential benefits and challenges.

²⁰ (Livingston, 2004)

²¹ For an example of a software package that may be used, See: <https://assess.com/xcalibre/>

Item response theory (IRT) attempts to mathematically model the probability of an identified response based upon the interaction of a person's latent ability and characteristics (e.g., difficulty and discrimination) of an item. IRT assumes that the latent construct being assessed and the items used to measure it are organised in an (unobservable) continuum. It focuses on establishing the test-taker's position on that continuum. This is calculated for each item independently, rather than focusing on the test as whole.

The rationale for using IRT in education settings is overcoming classical test theory's limitations in terms of item-level modelling and it can often be more accurate. The more advanced IRT models (2 or 3 parameters) are used to account for tests where discrimination and guessing are key challenges respectively.

IRT models, including the more basic, introductory Rasch model, do require at least one form of highly specialised software.

Significant psychometric expertise would also need to be in place to support building, maintaining, running, checking and analysing the model. It is essential that a team of psychometricians be fully employed on this work so that work can be cross-checked.

Due to the wide range of possible methodologies here, it would also be important to run this analysis centrally, rather than allowing individual models to be built by different teams involved in the

process; the risk would be inconsistency, as there is at present.

These complex models can often lack transparency, as it is challenging to explain their working to non-technical stakeholders. This risks a 'black box' effect, where the mechanics of how scores are calculated is opaque.

Overall

Depending upon the specific data sets, it may be worth experimenting with two equating methods and comparing the results before deciding which to use. The relative advantages of one method may best suit equating upper and lower test score tails. Combined with a (slightly) different equating method that is best applied to the main, central data points.²²

Recommendation 6:

It is probably most realistic to start with a more straightforward method, in terms of transparency, communicability, and capacity; this could be reviewed over time.

An equipercentile chained equating model seems to present the best balance between accuracy and these other concerns.

²² For a paper that shows the frequency estimation equipercentile equating method may be comparable to the *comparable outcomes* approach adopted by Examination Boards in the

UK, see: <https://www.cambridgeassessment.org.uk/Images/182461-using-statistical-equating-for-standard-maintaining-in-gcses-and-a-levels.pdf>

5. Scaling and reporting

The tests can be equated and the marks awarded scaled accordingly to be equivalent between test versions, thus taking account of any differences in overall test difficulty using the methods outlined above. While reporting and aggregation is beyond the explicit remit of the report, Cambridge offers the following as a suggestion of how the equated test scores could be used.

At test level, the scaled scores can be moved to a standardised scale. This allows direct comparisons to be made between test versions. Some aspects to be aware of include:

- The new scale needs to be distinctly different to avoid any confusion with the initial raw score scale. Test takers need to be fully aware that the raw score has been scaled. This avoids any confusion.
- Consideration should be given to how fine the scaled score should be. Each additional correct answer must clearly increase the test-takers' scaled score, but not by such a large difference that its importance is exaggerated.
- A decision will need to be made about whether to truncate the scaled scores. Truncating scaled scores means specifying a maximum or minimum value for the reported scaled scores that is less than the maximum or minimum value. This is sometimes considered for motivational purposes (e.g., to prevent test-takers from scoring '0' as a standardised score when they get an item correct).

Generally, a constant mean value would be set, and then force the scores into a normal distribution around this.²³ Just as the number of raw marks available should be constant between years, so should the standardised scale, as this will allow for direct comparisons between years to be made for monitoring purposes. Using a standardised scale will allow for constant, transparent comparisons across time, even if the test design changes.

The individual test scores within a subject would then need to be aggregated for each subject, taking into account any weightings that need to be applied. This will give a total standardised score for each test-taker in each subject.

This score will then need to be compressed into a 0-20 scale for the purposes of reporting each subject test, based on the scale currently in use.

It is important that this is the final stage, as doing this transformation prior to aggregation will result in a compounded loss of data precision.

Role of the national test scores in students' final outcomes

This score of 0-20 can be given to one decimal place, and then combined with a score derived from the teacher grade at the desired weighting.

It would be possible for the test to contribute more of the overall grade over time if this model is used. This overall output can then be assigned credits, as under the current system.

There may also be some merit in presenting both the teacher's grade for a student, expressed as a letter grade and based on the grading criteria as is

This would, however, require more human resource capacity.

²³ Alternatively, such a forced normal distribution could be avoided by using a linear transformation.

currently the case, in tandem with a numerical total mark, using a standardised scale as described above, for each individual subject.

In such a scenario, stakeholders would benefit from receiving a qualitative judgement as well as a more objective, performance-based measurement, and would then be free to use the dual results in such a way that they choose or as is most well suited to the context.

Implications for security and administration

Whichever methodology for standard setting is used, consideration must be given to the administration and security of the tests. No high stakes testing system can withstand regular and habitual compromises of test security. In any methodology, students who have advance sight of the assessment materials are inherently advantaged over candidates who have not. The results of any test that is not secure will therefore always be questioned.

The introduction of online test administration should help this, as it will be more difficult to get material out of the examination room, but care must be taken that candidates cannot access other internet-enabled services on the machine they are using to take the test.

Consideration might be given to:

- ad-hoc inspections
- computer-enabled malpractice spotting
- clearer guidance and consequences for leaking material
- use of AI-enabled systems for social media monitoring
- remote proctoring opportunities.

National tests for adult education

Where it is necessary to administer the test over a window of time, it is even more important that rules are clear and that discovery of breaking them has consequences for the test taker and those who enable them.

We understand that, at present, the national tests for adult learners are administered on-demand.

This causes many practical challenges in terms of test construction and crucially around the security of the tests. This may also threaten the validity of the results and affect perceptions of fairness around a process which leads directly to university admission and is therefore high stakes in nature.

Ideally adult learners would take the tests during a limited number of test windows as is the case for the upper secondary tests.

If on-demand testing continues to be used, then the test material used must be kept secure and not released to candidates in any form. In addition, it would be best if the test were one previously sat by a national cohort, so that the same thresholds/cut-scores can be automatically applied.

Bibliography and embedded links to documents

AERA (2014) 'Standards for Educational and Psychological Testing'

Baird, J., Isaacs, T., Opposs, D. and Gray, L. (Eds) (2018) 'Examination Standards: How measures and meanings differ around the world', UCL

CEM Centre (2008) 'Relative difficulty of examinations in different subjects', Durham University

Haladyna, T., Lane, S. & Raymond, M. (Eds) (2016), 'Handbook of Test Development' (2nd ed.), Routledge

Linacre, M, 'Anchored Estimation'. Available at <https://www.winsteps.com/winman/anchoredestimation.htm>

Livingston, S. (2004) 'Equating Test Scores (without IRT)', ETS

Ryan, J., Brockmann, F. (2011) 'A Practitioner's Introduction to Equating with Primers on Classical Test Theory and Item Response Theory', CCSSO

The national curriculum in England Key stages 3 and 4 framework:

[https://www.google.com/url?](https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&ved=2ahUKEwiXt5zDxLqJAXWYIhAIHSeRDRQQFnoECBUQAQ&url=https%3A%2F%2Fassets.publishing.service.gov.uk%2Fmedia%2F5da7291840f0b6598f806433%2Fsecondary_national_curriculum_corrected_PDF.pdf&usg=AOvVaw3DaffFvrKyj21Ls_DmViik&opi=89978449)

[sa=t&rct=j&q=&esrc=s&source=web&cd=&ved=2ahUKEwiXt5zDxLqJAXWYIhAIHSeRDRQQFnoECBUQAQ&url=https%3A%2F%2Fassets.publishing.service.gov.uk%2Fmedia%2F5da7291840f0b6598f806433%2Fsecondary_national_curriculum_corrected_PDF.pdf&usg=AOvVaw3DaffFvrKyj21Ls_DmViik&opi=89978449](https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&ved=2ahUKEwiXt5zDxLqJAXWYIhAIHSeRDRQQFnoECBUQAQ&url=https%3A%2F%2Fassets.publishing.service.gov.uk%2Fmedia%2F5da7291840f0b6598f806433%2Fsecondary_national_curriculum_corrected_PDF.pdf&usg=AOvVaw3DaffFvrKyj21Ls_DmViik&opi=89978449)

Curriculum for the compulsory school, preschool class and school-age educare:

[https://www.google.com/url?](https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&cad=rja&uact=8&ved=2ahUKEwj3rO3WxrqJAXmCRAIHeB8OU0QFnoECB4QAQ&url=https%3A%2F%2Fwww.skolverket.se%2Fdownload%2F18.31c292d516e7445866a218f%2F1576654682907%2Fpdf3984.pdf&usg=AOvVaw0xA0rMsbYabu3EpdwOy1-A&opi=89978449)

[sa=t&rct=j&q=&esrc=s&source=web&cd=&cad=rja&uact=8&ved=2ahUKEwj3rO3WxrqJAXmCRAIHeB8OU0QFnoECB4QAQ&url=https%3A%2F%2Fwww.skolverket.se%2Fdownload%2F18.31c292d516e7445866a218f%2F1576654682907%2Fpdf3984.pdf&usg=AOvVaw0xA0rMsbYabu3EpdwOy1-A&opi=89978449](https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&cad=rja&uact=8&ved=2ahUKEwj3rO3WxrqJAXmCRAIHeB8OU0QFnoECB4QAQ&url=https%3A%2F%2Fwww.skolverket.se%2Fdownload%2F18.31c292d516e7445866a218f%2F1576654682907%2Fpdf3984.pdf&usg=AOvVaw0xA0rMsbYabu3EpdwOy1-A&opi=89978449)

Cambridge Approach to Improving Education.

<https://www.cambridgeassessment.org.uk/Images/cambridge-approach-to-improving-education.pdf>



CAMBRIDGE

Contact us

The Triangle Building
Shaftesbury Road
Cambridge
CB2 8EA
United Kingdom

+44 (0) 1223 358331
partnership@cambridge.org

www.cambridge.org/partnership

[twitter.com/@CambridgePfE](https://twitter.com/CambridgePfE)

linkedin.com/showcase/cambridgepartnership

Transforming societies
through education

© Cambridge University Press & Assessment 2023



**Underlagsrapport till utredningen:
Likvärdiga betyg och meritvärden
(Dir.2023:95)**

UNDERLAGSRAPPORT

Likvärdiga betyg och meritvärden (Dir.2023:95)

Delområde: Den skarpa F-gränsens konsekvenser

Ida, Lindblad, Elisabeth Fernell, Christopher Gillberg*

*Ida Lindblad, leg. psykolog, med.dr., verksam i kommunala och fristående skolor och knuten till Gillbergcentrum, Sahlgrenska akademien, Göteborgs universitet, Kungsgatan 12, 411 19 Göteborg (ida.lindblad@gnc.gu.se)

Elisabeth Fernell, barnneurolog, professor i barn-och ungdomspsykiatri, Gillbergcentrum, Sahlgrenska akademien, Göteborgs universitet, Kungsgatan 12, 411 19 Göteborg, knuten till Neuropsykiatrisk mottagning barn och ungdom (BNK), Göteborg (elisabeth.fernell@gnc.gu.se)

Christopher Gillberg, barn-och ungdomspsykiater, professor i barn-och ungdomspsykiatri, Gillbergcentrum, Sahlgrenska akademien, Göteborgs universitet, Kungsgatan 12, 411 19 Göteborg, knuten till Neuropsykiatrisk mottagning barn och ungdom (BNK), Göteborg (christopher.gillberg@gnc.gu.se)

Introduktion

Utredningen om ”Likvärdiga betyg och meritvärden” (Dir. 2023:95) har avtalat om tre underlag från olika vetenskapliga discipliner vad gäller konsekvenser av den – i dagens betygssystem – skarpa gränsen för godkända betyg för tillträde till gymnasieutbildning.

Förfrågan till vår grupp har avsett en underlagsrapport till denna utredning, som ska gälla det aktuella kunskapsläget inom vår vetenskapliga disciplin; särskilt rörande hur elevers skolgång påverkas av att betygssystemet har en skarp gräns för godkänt vad gäller konsekvenser för tillträde till utbildning på gymnasial nivå.

Bakgrund och syfte

Vår forskargrupp träffar barn och ungdomar i klinisk sjukvårdsverksamhet och inom skolans elevhälsa i samband med bedömning och utredning av vad som kan ligga till grund för olika former av skolsvårigheter. Det kan handla om elever med svårigheter som rör inlärning, koncentration, kamratkontakt, språk, motorik och olika former av beteendemässiga svårigheter, t.ex. emotionell labilitet med benägenhet för utbrott. Vår uppgift som läkare och psykolog – i samverkan med pedagoger och för vissa barn också andra yrkeskompetenser – blir att kartlägga vilka funktioner barnet har svårt med, graden av svårigheter och vilka bakomliggande kognitiva funktioner som är involverade. Ett syfte med bedömningen är att kartlägga om svårigheterna stämmer med en, eller flera, funktionsnedsättningar, såsom exempelvis Attention-Deficit/Hyperactivity Disorder (ADHD), autism, intellektuell funktionsnedsättning (IF), eller om svårigheterna bedöms vara betydelsefulla, utan att kriterier för en funktionsbeskrivande diagnos uppfylls. I ett nästa steg tas ställning till bakomliggande medicinska faktorer och om en medicinsk orsaksdiagnos kan fastställas. Därefter ges information om den genomförda utredningen till vårdnadshavare och anpassat till eleven själv, samt till berörda inom skolan; pedagoger och elevhälsans professioner. Till sist diskuteras och planeras insatser och uppföljning. En sådan kartläggning och utredning startar inom elevhälsan och för många elever krävs fortsatt bedömning inom barnmedicinsk eller barnpsykiatrisk mottagning.

Vi har i våra arbeten särskilt berörts av de många elever som har svårigheter att uppnå godtagbara kunskapsmål i årskurs 3 och av de som från årskurs 6 inte når betygskriterierna för godkända betyg i ett eller flera ämnen. I många fall rapporterar deras lärare att barnet anstränger sig och gör sitt allra bästa för att lyckas i skolarbetet. Som ett barn (pojke 11 år med ADHD

och ”svag språklig begåvning”) själv uttryckte det: ”När alla andra är klara med uppgiften har jag ofta inte ens förstått vad vi ska göra”.

Vi har därför ställt oss frågan: Har alla barn i grundskolan förutsättningar att klara den nuvarande läroplanens krav? Det har aldrig undersökts vilka kognitiva förutsättningar som krävs för att uppnå lägsta godkända betyg (E).

En studie, som genomförts av vår forskargrupp omfattade elever i årskurs 6 (Lindblad m.fl. 2018). Den visade att elever som blivit underkända i minst ett ämne hade signifikant lägre kognitiva testresultat och fler lärarrapporterade svårigheter, t.ex. avseende koncentration, jämfört med barn utan underkända betyg. Den stora majoriteten av dessa barn med minst ett underkänt betyg hade fått stödsatser i skolan. Resultatet talar för att det finns elever i grundskolan som inte har förutsättningar att klara av de krav som läroplanen ställer.

Ett särskilt problem med betyget F uppstår i årskurs 9, då godkända betyg krävs bl.a. i matematik, engelska och svenska för att komma in på ett gymnasieprogram, inklusive på ett yrkesinriktat program. Vi genomförde därför en studie för att undersöka hur det går för elever som antingen inte kommit in på ett gymnasieprogram eller inte avslutat sina gymnasiestudier. Denna grupp jämfördes med elever som avslutat gymnasiet efter tre års studier. Undersökningen gjordes som en registerstudie och omfattade fem kommuner med elever födda 1998, som avslutat årskurs 9 år 2014 och som följts upp under år 2020 (Landgren m.fl. 2024). Jämfört med den grupp som kunnat fullfölja tre års gymnasiestudier hade de som inte kommit in på ett gymnasieprogram eller som inte avslutat sina gymnasiestudier en signifikant högre andel med arbetslöshet, psykosociala problem, inklusive kriminalitet, och psykiatriska diagnoser, såsom ADHD. Vår studie gav indirekt starkt stöd för att det behöver införas ett betygssystem som tillåter utbildning på ett yrkesprogram utan de teoretiska betygskrav som gäller i dag.

Syftet med denna rapport är att med stöd i vår och andras neuropsykiatriska forskning diskutera effekterna av dagens skarpa gräns för godkänt i det svenska betygssystemet.

Elevers kognitiva förmågor kan begränsa möjligheten att uppnå godkända betyg

Kognitiv utveckling och kognitiva funktioner

Den kognitiva utvecklingen pågår långt upp i åren. Vissa förmågor är inte fullt utvecklade förrän vid 20 till 25 års ålder. Kognitiv utveckling definieras i Nationalencyklopedin som ”utveckling av de intellektuella funktionerna i vid mening: tänkande, förståelse, beslutsfattande, memorerande, problemlösande, tolkande, jämförande, bedömande etc.” Den kognitiva utvecklingen kommer att påverka vilka förutsättningar barnet har att lära sig, både i skolan och i livet. Beroende på vilka kognitiva förmågor som är begränsade så kommer olika färdigheter att påverkas. Exempel på kognitiva förmågor är intellektuella, exekutiva (inklusive planering, uppmärksamhet och koncentration, arbetsminne, flexibilitet, motorisk planering) och sociala (inklusive social intuition).

Kognitiva begränsningar innebär att många av lärandeförmågorna påverkas såsom att tänka logiskt, dra slutsatser, analysera och ställa samman information, minnas och återge samt att på olika sätt bearbeta information. Kognitiva begränsningar av olika slag ingår vid alla neuropsykiatriska/utvecklingsneurologiska funktionsnedsättningar.

Utredning av kognitiva funktioner och funktionsnedsättning/ar och rättigheter i skolsystemet

Om ett barn har kognitiva begränsningar av sådan grad att det väcker oro och försvårar barnets funktionsförmåga i skolan så behöver detta bedömas genom utredning av psykolog och läkare, i samarbete med pedagog. För vissa barn behövs medverkan av ytterligare professioner, som logoped, fysioterapeut och arbetsterapeut. En sådan utredning kan påvisa vilka kognitiva grundsvårigheter barnet har och relateras till barnets funktionssvårigheter. Barnets symtom och svårigheter kan visa sig stämma med en eller flera funktionsnedsättningar, som till exempel intellektuell funktionsnedsättning, ADHD, autism, motorisk koordinationsstörning (Developmental Coordination Disorder, DCD) eller språkstörning. Det finns också barn med svårigheter inom olika kognitiva områden utan att symtombilden är av en grad som motsvarar kriterier för en funktionsnedsättning/diagnos. Det kan t.ex. handla om barn som har svag teoretisk begåvning inom ”normalvariationens” nedre del, eller barn som har autistiska drag, ”ADHD-drag” eller kombinationer av dessa.

Intellektuell funktionsnedsättning

För barn med IF (graderad enligt lindrig, medelsvår, svår och mycket svår form) finns sedan 1968 en särskild skolform, särskolan. Den bytte under 2023 namn till anpassade grundskolan, resp. anpassade gymnasieskolan. Den *största* gruppen elever i anpassade skolformer utgörs av barn med *lindrig* IF. För dessa barn finns en lagstadgad rättighet att få en till funktionsnedsättningen anpassad undervisning. Att bli mottagen i någon av dessa anpassade skolformer kräver att diagnos fastställts utifrån en kognitiv bedömning (psykologutredning) och en läkarbedömning, som behöver innefatta medicinska ställningstaganden till bakomliggande medicinsk orsak. Själva mottagandet i dessa anpassade skolformer kräver också utlåtande från pedagog och från kurator/socionom (Skolverket 2024).

Andra funktionsnedsättningar med rättigheter i skolsystemet

Vissa andra elevgrupper har också rätt till särskilda skolformer/”specialskola”, t.ex. barn med grav språkstörning, barn med grav hörselnedsättning eller dövhet och barn med synnedsättning i kombination med annan funktionsnedsättning. Läsåret 2022/23 gick 784 elever i specialskolan (årskurs 1–10) (Skolverket 2023).

Elever med grav språkstörning (impressivt och expressivt) är den största av dessa grupper. Antalet elevplatser i specialskolan för barn med språkstörning är sannolikt för lågt med hänsyn till behovet. Många föräldrar överklagar efter avslag från Specialpedagogiska skolmyndigheten (SPSM) till Skolväsendets överklagandenämnd. Två till tre procent av alla barn beräknas ha en svår/grav språkstörning. Majoriteten av dessa har ytterligare funktionsnedsättningar, oftast ADHD. Sannolikt har hälften av eleverna med svår/grav språkstörning behov av en anpassad skolform i specialskolan eller i språkklass (en mindre undervisningsgrupp i den ordinarie grundskolan). Riksförbundet för döva, hörselskadade barn och barn med språkstörning (DHB) har betonat behovet av bättre stöd för de barn som har tydliga språksvårigheter men som inte anses kvalificerade att ingå i specialskolans målgrupp (Miniscalco m.fl. 2018, Stübner m.fl. 2020).

Utbildning för rörelsehindrade ungdomar (Rh-anpassad utbildning)

Ungdomar med svåra rörelsehinder har ofta behov av för dem anpassad utbildning. Sådan utbildning avser gymnasieutbildning (på något av riksgymnasierna) för ungdomar med svårt rörelsehinder, t.ex. till följd av cerebral pares, svår muskelsjukdom eller annan neurologisk

sjukdom (Skolverket 2024). Särskild skolform finns inte längre för dessa elever inom grundskolan. Alla elever som söker Rh-anpassad gymnasieutbildning tas ej emot. Bland annat har studier från Göteborg visat att elever med cerebral pares i hög utsträckning, i kombination med rörelsenedsättning, har kognitiva svårigheter i form av autism och/eller ADHD (Påhlman m.fl. 2021). Dessa elever skulle gynnas av sådan Rh-anpassad skolform med mer individuellt stöd, även om rörelsenedsättningen i enskilda fall inte är av den betydande svårighetsgrad som anges som krav för tillträde i skollagen.

Svag teoretisk begåvning

Jämfört med lindrig IF, som förekommer hos ca 1,5 procent av barnpopulationen, uppgår gruppen barn med svag teoretisk begåvning, inom den vida normalvariationen, dvs. en svagare generell/intellektuell förmåga, till 12–14 procent (IQ-område ca 70 till 84). Dessa elever har ofta även någon grad av svårigheter avseende exekutiva förmågor. För denna elevgrupp finns inte någon given anpassad undervisning, flertalet har inte heller fått sina svårigheter identifierade genom en kognitiv bedömning/testning av psykolog. Till exempel visade vi i en studie av elever i årskurs 6 (Lindblad m.fl. 2018) att elever som fått underkänt betyg i minst ett ämne hade signifikant lägre kognitiva resultat i psykologtest (WISC-IV) (Wechsler Intelligence Scales for Children, version 4) och fler lärarrapporterade svårigheter jämfört med barn utan underkända betyg. Studien visade att 30 procent av eleverna med svaga resultat på kognitiva test hade fått underkänt betyg i minst ett ämne, jämfört med tre procent av eleverna med ett genomsnittligt kognitivt testresultat. Studien visade således att barn med svaga kognitiva testresultat i mycket högre utsträckning riskerar att inte nå godkända betyg (A–E) i alla ämnen i grundskolan. Samtliga, utom en elev, i vår studie hade fått stödinsatser.

Dyslexi

I en studie av barn med dyslexi visades att det förelåg en mer än åttafaldig ökning av olika kognitiva funktionsproblem och funktionsnedsättningar i denna grupp jämfört med i en grupp utan läs- och skrivsvårigheter (Brimo m.fl. 2021).

Särskilda brister avseende exekutiva funktioner

I läroplanen, Lgr 11 och även Lgr 22, ställs i dag särskilt stora krav på barns exekutiva förmågor, dvs. att planera, organisera, reflektera och självständigt utföra och slutföra skolarbete. Brister i exekutiva förmågor föreligger förutom vid ADHD också vid andra kognitiva funktionsproblem och funktionsnedsättningar, inklusive autism. Detta medför att

barn med svag teoretisk begåvning, barn med koncentrationssvårigheter, med språkliga svårigheter och med svårigheter avseende socialt samspel får problem att klara ett eller flera av de betygskriterier som krävs för godkänt betyg.

I boken *Kunskapssynen och pedagogiken* (Enkvist m.fl. 2017) belyses att aktuell läroplan vid betygssättning – mer än inläring och faktakunskaper – betonar elevers förmåga att föra resonemang och diskutera, dvs. ställer större krav på exekutiva förmågor än på enklare kunskaper och färdigheter i olika ämnen.

Elevers kognitiva förutsättningar måste beaktas i läroplan, undervisning och vid betygssättning

Hur många i elevgruppen har kognitiva funktionsproblem/funktionsnedsättningar?

Baserat på prevalenser, dvs. förekomst av olika funktionsnedsättningar och med hänsyn tagen till att det praktiskt taget alltid är fråga om kombinationer av svårigheter, kan den totala andelen av elever med stora behov av anpassningar anges till ca 10 procent av den totala elevgruppen (ca 13 procent av pojkarna och sju procent av flickorna). Begreppet ESSENCE (Early Symptomatic Syndromes Eliciting Neurodevelopmental Clinical Examinations) formulerades av Gillberg (2010) och belyser den samvariation som finns mellan olika kognitiva funktionsnedsättningar och funktionsproblem såsom ADHD, autism, intellektuell funktionsnedsättning, språkstörning, Tourettes syndrom, motoriska koordinationsvårigheter och svag teoretisk begåvning (Landgren m.fl. 2024a). ”Helhetssyn är grunden inom ESSENCE, ett samlingsbegrepp för tidigt debuterande utvecklingsneurologiska problem – problem som till stor del överlappar varandra” (Gillberg 2014).

Alla elevers olika förutsättningar måste beaktas

I grundskolans skolmiljö förväntas elever klara av att planera och organisera sitt lärande, hålla uppmärksamheten, styra sina impulser, förstå tid och vara en fungerande del i det sociala sammanhanget. När de kognitiva förmågorna är begränsade påverkas barnets fungerande i skolan och för många även möjligheten att nå godkända betyg.

Vi har således en skolform för elever med intellektuell funktionsnedsättning, en grupp som uppgår till ca två procent av populationen. Den anpassade grundskolan och den anpassade gymnasieskolan erbjuder dessa elever en läroplan som är anpassad till deras kognitiva förutsättningar.

För elever med svag teoretisk begåvning, även för den grupp som ligger strax ovan gränsen för lindrig intellektuell funktionsnedsättning, finns dock ingen given anpassad skolform. Dessa elever har dessutom ofta andra svårigheter eller funktionsnedsättningar och vissa uppfyller kriterier för ADHD. En del elever med svag teoretisk begåvning har samtidigt en språkstörning, en del har autism eller autistiska drag. Motorisk koordinationsstörning är också vanligt förekommande. Inget av detta beaktas specifikt i någon befintlig skolform utan huruvida dessa elever får sina behov tillgodosedda är avhängigt huvudmannens prioriteringar och enskildas goda vilja.

Vi bedömer att elever med svårigheter i skolarbetet bör få en första kartläggning inom elevhälsan. Kartläggning bör ske i samverkan mellan pedagoger, psykolog och skolläkare. Då kan det framkomma om eleven har en svag teoretisk begåvning och/eller symtom som talar för autism, ADHD eller annan specifik problematik. Därefter kan man ta ställning till om eleven behöver ytterligare utredning.

Vid konstruktion av ett betygssystem inklusive formulering av betygs-kriterier måste hänsyn tas till elevers olika kognitiva förutsättningar. Elever med svag teoretisk begåvning i kombination med betydande exekutiva svårigheter löper mycket stor risk att inte uppnå godkänt betyg i ett eller flera ämnen. Det skolsystem som beskrivs i skollagen (2010) och i Lgr 22 (Skolverket 2022), dvs. en skola som ska ”främja alla elevers utveckling och lärande samt en livslång lust att lära” – är fortfarande ett skolsystem att önska att döma av de olikvärdiga möjligheter och förutsättningar som elever ges. Det nuvarande betygssystemet riskerar också att påverka elevers självkänsla negativt. Många pedagoger har, som Arne Engström, professor i matematikdidaktik, framfört att ”ett underkänt betyg inte borde höra hemma i en obligatorisk skola” (Svenska Läkaresällskapet 2021).

En skolas uppgift måste vara att säkerställa att alla barn får utvecklas och lyckas efter sina förutsättningar. Barns naturligt begränsade kognitiva förmågor – med hänsyn till ålder och mognad – kommer att påverka lärandet, och i läroplanens krav på kognitiva/exekutiva förmågor måste detta beaktas. Grundläggande ämneskunskaper och färdigheter har minskat i värde till förmån för att resonera, reflektera och ifrågasätta – utifrån olika perspektiv (Enkvist m.fl. 2017; Skogstad 2019). Det behöver vara möjligt att delta i undervisningen och uppnå de grundläggande målen i läroplanen med dess kurs- och ämnesplaner utan att kunna resonera om innehåll i olika kontexter på ett alltför abstrakt plan. Målet måste vara att kursplanen/läroplanen

utföras så att alla elever har möjlighet att nå åtminstone de kunskaper som krävs för godkända betyg.

Det behövs fler möjligheter för elever med autism

Fram till 2011 hade elever med autism i kombination med en svag teoretisk begåvning, nära gränsen för lindrig IF, rätt att få sin undervisning inom särskolan om detta bedömdes gagna elevens inläring och skolfungerande i stort. Efter att Lgr 11 trädde i kraft blev dock särskolan, och senare den anpassade grundskolan, begränsad till att strikt omfatta barn med IF. I detta sammanhang kan nämnas den grupp av förskolebarn som har diagnostiserad autism och en generell utvecklingsförsening – med betydande språklig funktionsnedsättning – som vid tiden för utredning fått diagnosen global utvecklingsförsening. Dessa barn har betydande kognitiva svårigheter, men intellektuell funktionsnedsättning har inte med säkerhet kunnat fastställas, och de får då börja i ordinarie förskoleklass trots uppenbara risker för att de inte kommer att kunna utveckla de kunskaper som krävs för godkända betyg. För denna elevgrupp bör möjligheten att få starta skolgången i den anpassade grundskolan därför återigen övervägas.

Elever med autism som ej har IF behöver också särskilda anpassningar i skolan. Elever med autism och teoretisk begåvning inom ”normalvariationen” har specifika språkliga/exekutiva svårigheter, innefattande brister i pragmatisk förmåga, dvs. i att använda språket i socialt samspel. Carlsson m.fl. (2020) visade att dessa elever hade svårigheter med språklig bearbetning och att klara uppgifter som innebar att berätta. För barn utan IF, men med svår autism behövs individuella lösningar i fråga om hur undervisningen skall anpassas.

Anpassade undervisningsgrupper

Historiskt har Sverige haft undervisningsgrupper för elever med svårigheter att klara kraven i den nationella läroplanen. Fram till 1980 fanns en särskilt anpassad undervisningsform som kallades hjälpklass. Där kunde elever med en svagare teoretisk begåvning, men inom ”normalvariationen”, få extra stöd i undervisningen. I dag finns i vissa skolområden s.k. resursklasser och särskilda undervisningsgrupper. Med tanke på behovet är det dock en otillräcklig dimensionering av mindre undervisningsgrupper.

Ett skolsystem måste ta hänsyn till hela elevgruppen, elevens olika kognitiva förutsättningar och behov av anpassningar. Det måste vara möjligt att anpassa skolan så att inte 14–15 procent av eleverna behöver klassas som obehöriga till gymnasiestudier (Skolverket 2023) och så att

inte den psykiska ohälsan till följd av skolstress ökar vid varje ny mätning som Folkhälsomyndigheten gör (2023). Det handlar om att anpassa läroplanen efter barns kognitiva förutsättningar vid olika åldrar och dessutom om att ta hänsyn till de elevgrupper som har specifika kognitiva funktionsbegränsningar.

De mest grundläggande målen bör omfatta färdigheter som alla elever har förutsättningar att utveckla och uppnå.

Klasstorlek

I en studie från Institutet för arbetsmarknads- och utbildningspolitisk utvärdering (IFAU 2012) anges att endast ett fåtal tidigare studier undersökt om det finns långsiktiga effekter av klasstorleksminskningar. I en annan studie från IFAU (Fredriksson m.fl. 2012) följde man de långsiktiga effekterna av förändringar av klasstorlekar på mellanstadiet. Författarna sammanfattar:

Vi studerar de långsiktiga effekterna av klasstorleksförändringar i årskurs 4 till 6. Mindre klasser förbättrar elevernas kognitiva och icke-kognitiva förmågor vid 13 års ålder, resultat på nationella prov i svenska, engelska och matematik vid 16 års ålder samt utbildningsnivå och löner som mäts vid 27–42 års ålder. Löneeffekterna är tillräckligt stora för att intäkterna ska överstiga de direkta kostnaderna av klasstorleksminskningen (s.1).

Det finns alltså all anledning att dimensionera klassernas storlek med hänsyn till det stöd som eleverna behöver så att lärare ges möjlighet att hjälpa de elever som behöver ett extra stöd.

Exempel på hur bör pedagogiken bör anpassas

För att barn med begränsade kognitiva förmågor ska kunna ta till sig undervisning bör de få undervisning som är utformad så att den innehåller mer av praktisk övning och tillämpning. Det innebär att kunskap som ska läras behöver sättas i ett sammanhang och tränas utifrån detta sammanhang på ett mer konkret sätt, i stället för att som nu sker, teoretisera kunskaperna. Studier inom kognitiv neurovetenskap visar att man genom att konkretisera undervisningens innehåll, och att få barn att uppleva det som de ska lära sig, ökar möjligheter till lärande för barn generellt, men särskilt för barn med svag teoretisk begåvning och/eller exekutiva begränsningar. Att automatisera kunskaper leder på sikt till ökade möjligheter att lära sig ”högre” kunskaper, dvs. att lära sig att reflektera och analysera. Inom matematiken finns ett starkt samband mellan automatisering av tabellkunskaper (multiplikationstabellen) och resultat

på de nationella proven i årskurs 6 och årskurs 9. En positiv korrelation kunde visas mellan automatisering av kunskap och problemlösningsförmåga (Joelsson 2015).

Grundläggande krav bör bygga på mer automatiserade kunskaper och mindre på exekutiva förmågor

Exekutiva förmågor

Elevens varierade kognitiva/exekutiva förmågor måste tas hänsyn till i undervisningen, och de som har svagare sådana funktioner måste uppmärksammas och ges pedagogiska stödinsatser. Nämda förmågor ska inte i sig vara betygsgrundande men de påverkar elevens möjlighet att lyckas i undervisningssituationer. I funktionsnedsättningen ADHD ingår som ”huvudsaklig svårighet” brister i de exekutiva funktionerna. Dessa innefattar förmågor såsom planering, organisering, arbetsminne, förmåga till flexibilitet, motivation och emotionell reglering. De ökade kraven på att elever ska ta eget ansvar för sitt lärande ställer särskilda krav på exekutiva förmågor, och därmed riskerar det att bli de exekutiva förmågorna som betygssätts och inte ämneskunskaperna.

I lärandet har de exekutiva funktionerna en viktig roll för elevers möjligheter att kunna tillgodogöra sig undervisningen. Detta innebär att elevens planeringsförmåga och studiestrategier behöver uppmärksammas och stöttas i undervisningen. Det gäller elever som i många situationer behöver få ett ”stöd för tanken”, en ”starthjälp”, en ledtråd, en påminnelse. De ökade klasstorlekarna försvårar lärarnas möjligheter att hjälpa enskilda elever med sådana ”små” insatser. Dessutom är det i dag inte en självklarhet att det i huvudsak är lugnt i klassrummet. Att elever får hörselkåpor för att kunna avskärma sig i klassrummet är ett inte ovanligt inslag. Detta trots att skolan ska betraktas som elevernas arbetsmiljö och att Arbetsmiljöverket konstaterar att hörselskydd ska vara en sistahandsåtgärd när det inte är möjligt att förebygga uppkomst av buller. Hög ljudnivå i klassrummet försämrar möjligheten till koncentration och är särskilt negativt för elever med exekutiva svårigheter.

För att nå minst E betyget ställer läroplanen krav på förmågor som ”att föra resonemang om”, ”underbyggda resonemang om”, ”formulera ställningstaganden med enkla motiveringar”, i stället för att enbart precisera vilka grundkunskaper som eleverna förväntas ha efter vissa årskurser. För många elever hade det varit enklare om det vore tillräckligt att minnas fakta, återge givna förklaringar, genomföra en uträkning enligt en given metod eller formulera en fras enligt en instruktion.

De exekutiva förmågorna är inte fullt utvecklade hos barn i grundskolan, utan dessa utvecklas under hela uppväxten. Detta innebär att skolans läroplan måste vara anpassad till detta förhållande. Kraven i gällande läroplan har aldrig, vad vi vet, satts i relation till vad som kan förväntas vad gäller kognitiv/exekutiv förmåga hos barn i allmänhet i olika åldrar. Läroplanens mål bygger snarare på de krav samhället ställer, vilka inte alla har möjlighet att leva upp till. Den nuvarande läroplanen som ställer krav på att kunskap – för att kunna betygssättas – ska kunna reflekteras kring eller resoneras om, ställer höga krav på de exekutiva förmågorna. Måluppfyllelsen blir särskilt svår för de grupper av elever som har svaga exekutiva förmågor – elever med koncentrationssvårigheter, elever med svag intellektuell förmåga, elever med autism och för dem med språkstörning.

Vikten av att automatisera kunskaper för att avlasta arbetsminnet

Genom att låta barn automatisera ämneskunskaper, genom träning av sådana kunskaper, ökar möjligheterna för barn att vara i lärande och att få uppleva att de lär sig. Det är i linje med Lgr 22: ”Varje elev har rätt att i skolan få utvecklas, känna växandets glädje och få erfara den tillfredsställelse som det ger att göra framsteg och övervinna svårigheter”. Genom att träna på att minnas fakta, förklaringar och rutinmetoder, ökar möjligheten att dessa lagras, då det i lägre utsträckning ställs krav på de exekutiva förmågorna vid denna typ av inläring. Genom att lära elever sådana mer grundläggande ämneskunskaper ökar möjligheterna för barn att senare i livet kunna bearbeta redan inlärd kunskaper genom att reflektera och resonera kring dessa.

Detta utgår från den kognitionsforskning som visar att arbetsminne och uppmärksamhetsförmåga är särskilt nedsatta hos individer som har bristande exekutiva förmågor (t.ex. barn med ADHD, autism eller svag teoretisk begåvning). För att barn ska ha möjlighet att utveckla lärandet krävs att kunskaper automatiseras (Whitehurst 2003). Inom matematikundervisningen har Bentley (2009) visat att om arbetsminnet kan hämta resultat av aritmetiska beräkningar (redan automatiserad kunskap), t.ex. multiplikationstabellen, från långtidsminnet så kan mer komplicerade beräkningar utföras. Om sådan automatiserad kunskap är otillräcklig kommer arbetsminnet att vara fullt upptaget med att utföra beräkningar som skulle kunna vara automatiserade; lärande uteblir när hjärnan upptas av dessa basala beräkningar i stället för att ägna sig åt mer komplicerade uppgifter. Detta får då konsekvenser för den matematiska problemlösningen och för lärandet generellt. Även Vygotsky (1978) framhöll att elevernas problemlösningsförmåga effektiviseras i takt med deras förmåga att

snabbt plocka fram redan inlärd/automatiserad fakta. Automatisering av kunskaper lönar sig i längden och bidrar till ökade kunskaper. Studier inom matematik visar också att högpresterande elever anstränger sig mycket mindre än vad lågpresterande gör, eftersom de använder sina minnesfunktioner på ett mycket mer effektivt sätt (Byers och Erlwanger 1985). Martin Ingvar (2009) visar att ny inläring underlättas av automatiserad kunskap eftersom all ny kunskap framkallas med hjälp av information som redan är lagrad i långtidsminnet. Det innebär att ju mer som finns lagrat i minnet desto snabbare går informationsbearbetningen. Vidare menar Ingvar att en av hjärnbarkens mest grundläggande funktioner är att all inläring ska leda till automatisering för att spara energi i lärandet. Slutsatsen är alltså att automatiserade kunskaper inte belastar arbetsminnet som är begränsat, särskilt hos yngre barn och hos barn med bristande exekutiva förmågor. Tidig kunskapsautomatisering får således på sikt positiva konsekvenser för betyg och meritvärden. Barn kommer att lära sig mer och på en högre nivå om de först ägnar tid åt att automatisera kunskaper i stället för att tidigt föra teoretiserande resonemang som belastar de exekutiva förmågorna.

De uppgifter som eleverna får i skolan i dag är ofta språkliga i sin utformning. Elever kan t.ex. använda olika text och bildkällor för att konstruera egna skriftliga produktioner eller muntliga framföranden om ett ställningstagande eller lösning på ett problem. Det skapar stora utmaningar för vissa elever. I de lägre årskurserna kan detta illustreras med följande skildring som illustrerar att konkretisering underlättar inläring och möjlighet att befästa kunskaper:

En lärare berättade att för 20 år sedan gick hon ut i skogen en hel eftermiddag och plockade svamp med sina elever. Eleverna fick bestämma art, lukta på svamparna och lära sig se skillnad på kremla och sopp. För tio år sedan fanns inte tiden längre, så läraren effektiviserade och tog med sig svampen in i klassrummet. Eleverna fick turas om att titta, lukta och artbestämma de svampar som låg vid katedern. I dag finns inte tid ens för detta, så nu får eleverna i stället titta på filmer om svampar, läsa en faktatext och skriva begreppen på ett uppgiftspapper. Det är så mål- och resultatstyrningen påverkar hur vi undervisar (Fohlin 2024).

I betygssättningen behöver barn enligt gällande läroplan visa sina kunskaper vid olika slags provsituationer. Provsituationerna är inte alltid anpassade till elevers kognitiva förutsättningar då de ska tjäna till att visa att eleven lärt sig att föra resonemang om, reflektera kring och motivera sina ställningstaganden. Sådana provsituationer ställer återigen höga krav på de exekutiva förmågorna. Skillnaden mellan ett prov där exekutiva förmågor ska visas och ett prov där ämneskunskaper ska redovisas är mycket stor. I en provsituation där eleven ska visa

ämneskunskaper kommer krav att ställas på barnets förmåga att hämta lagrad information, och i denna process krävs minimal exekutiv förmåga. Har dessutom ämneskunskapen automatiserats kan man se det som att en ”motorväg”, tar kunskapen ”förbi” de exekutiva funktionerna och direkt till de kognitiva delar av hjärnan där kunskapen lagras. Om lärare ges möjlighet att öva ämneskunskaper med eleverna ökar möjligheten för dessa ”motorvägar” att utvecklas i hjärnan.

Starka skäl att ompröva det rådande betygssystemet som exkluderar elever från vidare studier

I Skolverkets rapport (2019) skrivs att ”Förskolan och skolan ska verka för att barns och elevers erfarenheter av lärandet och undervisningen också främjar deras hälsa och välbefinnande”. Samtidigt rapporterar Folkhälsomyndigheten (2023) att trivseln i skolan fortsätter att minska sedan 2013/2014 för både flickor och pojkar i åldrarna 13 och 15 år. Parallellt fortsätter skolstressen att öka sedan 2009/10 för både flickor och pojkar i samtliga undersökta åldrar (11, 13 och 15 år). Mot denna bakgrund bör det finnas skäl att ompröva den rådande läroplanen och det rådande betygssystemet, som elever, skola och pedagoger i dag måste förhålla sig till.

Skolverket rapporterar att ca 14–15 procent av en årskull elever inte når gymnasiebehörighet efter årskurs 9 (Skolverket 2023). Dessa elever har minst ett icke-godkänt betyg i något av ämnena matematik, svenska eller engelska. En motsvarande betygssättning finns inte i andra europeiska länder eftersom man i de flesta skolsystem inte underkänner elever inom den obligatoriska skolan och därmed förhindrar dem från tillträde till fortsatt utbildning (Ågren och Bremberg 2022).

Uppföljning av en svensk elevgrupp som år 2014 avslutade årskurs 9 visar att av de elever som inte kommit in på ett nationellt gymnasieprogram – vid uppföljning sex år efter avslutad grundskola – hade en betydligt högre andel neuropsykiatriska funktionsnedsättningar, var arbetslösa, behövde olika former av sociala stödinsatser och var i högre grad kriminellt belastade – jämfört med den grupp som hade påbörjat och kunnat avsluta tre års gymnasiestudier (Landgren m.fl. 2024b).

Situationen i Sverige bör kunna knytas till den reformering av gymnasieskolan som skedde 1991 (prop. 1990/91:85). Då fattades beslut om att införa en ny gymnasieskola där alla

program, såväl yrkesförberedande som teoretiska, skulle vara treåriga och kvalificera för högskolestudier (Erikson 2018; Ågren och Bremberg 2022). Sverige fick därmed en av världens mest sammanhållna gymnasieskolor där samtliga elever förväntades få grundläggande behörighet till universitetet. Före 1991 var yrkesutbildningen skild från den teoretiskt inriktade gymnasieskolan. År 1991 inordnades således yrkesutbildningen i den teoretiska gymnasieskolan och krav ställdes på ökade teoretiska kunskaper i kärnämnen svenska, engelska och matematik.

Sverige hör nu till de länder i Västeuropa som har den högsta dödligheten bland unga vuxna i åldern 20–34 år. Enligt ”Global Burden of Disease” (Institute for Health Metrics and Evaluation 2024), började dödligheten i Sverige avvika från det västeuropeiska genomsnittet för åldersgruppen 20–24 år omkring 1996, för åldersgruppen 25–29 år omkring år 1999 och för åldersgruppen 30–39 år omkring år 2007. Åldersgruppen 20–24 år i studien skulle förväntats ha börjat gymnasiet, om de hade behörighet, vid 16 års ålder. Om de hade påverkats negativt av reformen 1991, skulle effekterna beräknats starta 1996. Det var också vad som visades (Bremberg 2024). Den enskilt viktigaste dödsorsaken i dessa åldrar i Sverige är självmord. I Västeuropa som helhet har självmorden sjunkit medan de i Sverige i dessa åldrar varit oförändrade. En svensk undersökning visade att risken för att begå självmord var sex gånger högre för unga med ofullständiga betyg från grundskolan jämfört med unga med goda betyg (Björkenstam m.fl. 2011).

Tjugo år senare, i Lgr 11 (2011), ökade kraven på teoretiska kunskaper ytterligare och för behörighet till ett yrkesprogram krävdes godkända betyg i kärnämnen svenska, engelska och matematik samt i ytterligare fem ämnen.

Viktigt att beakta barns olika förutsättningar vad gäller kognitiva förmågor

Kursplaner måste ta hänsyn till att grundskolan arbetar med barn under utveckling av de exekutiva förmågorna (som t.ex. arbetsminne och uppmärksamhet) och att en betydande del av populationen (12–14 procent) har en svag teoretisk begåvning. Därför måste realistiska lärandekrav ställas på dessa ännu utvecklade förmågor och lärandet måste anpassas till barn och ungdomars förutsättningar vad gäller förmåga att automatisera grundläggande kunskaper. När dessa förutsättningar har skapats måste också betygssättningen utformas på ett sätt som

gör att barn upplever det möjligt att nå realistiskt satta gränser för godkända betyg så att deras motivation kan bibehållas.

Några resultat från forskning om betyg i Sverige

Sveriges kommuner fick själva bestämma om det skulle ges betyg i årskurs 6 mellan åren 1969 och 1981. Detta har gjort det möjligt att undersöka hur tidiga betyg påverkat eleverna. Skolverket (2020) har tagit upp att införandet av betyg i årskurs 6 har lett till ökad stress och försämrad akademisk självkänsla kopplad till skolan hos elever, och man refererar bl.a. till studier av Högberg m.fl. (2021). Flera andra har också funnit att beslutet att införa betyg i skolan i årskurs 6 riskerar att få den omvända effekten för elever som har svårigheter i lärandet (Boström och Bostedt 2020). Forskning från Örebro universitet (Lundahl m.fl. 2015) har givit stöd för att betyg generellt differentierar och påverkar äldre och yngre elever och låg- och högpresterande elever på olika sätt. Lågpresterande och yngre elever tycks påverkas mer negativt av betygsättning jämfört med äldre och högpresterande. Liknande slutsatser – att betyg påverkar elever på olika sätt – rapporterades av Klapp (2008, 2015), som fann att högpresterande elever varken påverkades negativt eller positivt av betyg, medan lågpresterande elever påverkades negativt.

Dessa studier har framför allt behandlat betygsfrågan vad gäller yngre elever. En avgörande fråga är dock hur det existerande betygssystemet hanterar ett icke godkänt betyg (F) i ett kärnämne inför ansökan till ett gymnasieprogram. Som vi redan noterat finns starka skäl att revidera det rådande betygssystemet, som exkluderar elever med betyget F i ett kärnämne från vidare studier.

De svagaste eleverna ska ges så bra förutsättningar som möjligt att hitta en meningsfull och produktiv plats i samhället som vuxna

Vi konstaterar att en icke försumbar andel av eleverna i den totala populationen saknar, eller har ringa förutsättningar att uppnå, vad samhället har satt upp som en godtagbar nivå i form av godkända betyg för tillträde till gymnasial utbildning. En elev som med avseende på någon del av betygsriterierna för ämnet inte når upp till den godtagbara nivån får i dag betyget F, oavsett om eleven är långt ifrån att uppnå dessa eller om eleven ligger strax under godtagbar nivå.

Fler elevgrupper måste ges möjlighet till undervisning i för dem anpassade undervisningsformer, inte enbart elever med IF. Det bör återigen göras en översyn av

elevgruppen med autism och samtidig svag teoretisk begåvning, på gränsen mot lindrig IF. Denna elevgrupp bör ges möjlighet till en anpassad grundskola, eller en motsvarande skolform, för att möjliggöra en positiv utveckling eftersom dessa elever ofta redan från unga år saknar förutsättningar att följa grundskolans läroplan. Elever med mycket svag intellektuell förmåga, på gränsen mot lindrig IF, en grupp som mycket ofta även har andra svårigheter, såsom koncentrationssvårigheter, språkstörning och som kan ha svårigheter inom autismområdet, behöver också anpassade undervisningsformer.

Elever med mer praktisk än teoretisk begåvning måste få möjlighet att under gymnasietiden få en praktisk yrkesutbildning, som inte vid antagning kräver godkända betyg i matematik, svenska och engelska. Utbildningsväsendet måste erbjuda yrkesprogram, utan de behörighetskrav i teoretiska ämnen vid ansökan, som de högskoleförberedande programmen ställer. Vi ser detta som ett av de allra viktigaste behoven eftersom det skulle innebära att alla elever i gymnasieåldern får en skoltillhörighet/undervisningssituation som ger dem möjlighet att få lyckas – och därmed hopp om ett framtida yrkesliv.

Sammanfattning

Skolsystemet måste när det gäller undervisning och betygssystem beakta den variation som föreligger i barns kognitiva förmågor; avseende den generella kognitiva/intellektuella förmågan, de exekutiva förmågorna och avseende förmågor som rör socialt fungerande.

Minst 10 procent av eleverna har svårigheter av en grad motsvarande funktionsnedsättning/ar inom dessa områden. Ca 12–14 procent av alla elever har en svag teoretisk begåvning, utan att uppfylla kriterier för intellektuell funktionsnedsättning. Svårigheter som dock ger problem med måluppfyllelse i dagens skola.

Skolsystemet måste beakta att inte endast elever med rättighet till den anpassade grundskolan, dvs. elever med intellektuell funktionsnedsättning, samt elevgrupper med rätt till undervisning i specialskolan, dvs. elever med grav språkstörning, med svår hörselnedsättning/dövhet och/eller med synnedsättning och ytterligare funktionsnedsättning, har behov av anpassningar/särskilda skolformer.

Kraven på att i skolarbetet självständigt planera, organisera, reflektera, analysera och resonera har ökat utan att dessa krav anpassats, varken till barns kognitiva utveckling generellt eller till

elever med specifika exekutiva svårigheter. Förekomsten av ADHD, vars huvudsårighet innebär kognitiva/exekutiva brister, är bland barn och unga 5–7 procent.

Det finns funktionssvårigheter och funktionsnedsättningar som medför att elever i grundskolan inte har möjlighet att nå de krav som är uppställda för godkänt betyg. Undervisning och betygskriterier måste anpassas till hela den elevgrupp som tillhör grundskolan. Av den totala elevgruppen har ca 10 procent minst en utvecklingsneurologisk/neuropsykiatrisk funktionsnedsättning. Kognitionsforskningen måste ligga till grund för pedagogiken. Det behöver i skolan bli ett ökat fokus på grundläggande kunskaper och färdigheter i ämnen och att dessa får prioritet i relation till exekutiva förmågor, som att reflektera och analysera. Slutbetyg i grundskolans årskurs 9 utgör en signal avseende de möjligheter som öppnar sig mot yrkeslivet för en ung människa. Betyget underkänt (F) bör inte användas i en obligatorisk skola.

Referenser

- Bentley, P.-O. (2009). TIMSS 2007 – En djupanalys av svenska elevers matematikkunskaper. *Nämaren*, 2009(1), 4–8.
- Björkenstam, C., Weitoft, G. R., Hjern, A., Nordström, P., Hallqvist, J., & Ljung, R. (2011). School grades, parental education, and suicide – A national register-based cohort study. *Journal of Epidemiology & Community Health*, 65(11), 993–998.
- Boström, L., & Bostedt, G. (2020). What about study motivation? Students' and teachers' perspectives on what affects study motivation. *International Journal of Learning, Teaching and Educational Research*, 19(5), 40–59.
- Bremberg, S. (2024). An educational reform in Sweden with unintended negative health effects. *Acta Paediatrica*. Advance online publication. <https://doi.org/10.1111/apa.17000>
- Brimo, K., Dinkler, L., Gillberg, C., Lichtenstein, P., Lundström, S., & Åsberg Johnels, J. (2021). The co-occurrence of neurodevelopmental problems in dyslexia. *Dyslexia*, 27(3), 277–293.
- Byers, V., & Erlwanger, S. (1985). Memory in mathematical understanding. *Educational Studies in Mathematics*, 16(3), 259–281.
- Carlsson, E., Åsberg Johnels, J., Gillberg, C., & Miniscalco, C. (2020). Narrative skills in primary school children with autism in relation to language and nonverbal temporal sequencing. *Journal of Psycholinguistic Research*, 49(4), 475–489.
- Enkvist, I., Henrekson, M., Ingvar, M., & Wällgren, I. (2017). *Kunskapssynen och pedagogiken* (M. Henrekson, red.). Stockholm: Dialogos Förlag.
- Engström, A. (2021). Eleverna behöver verkligt stöd – inte tomma kulisser. *Skola och Samhälle*. <https://www.skolaochsamhalle.se/flode/skola/arne-engstrom-cleverna-behover-verkligt-stod-inte-tomma-kulisser>
- Erikson, J. (2018). Gymnasireformen 1991 – en skola för jämlikhet eller en skola för arbetsmarknaden? *Vägval i skolans historia, årg. 2018, nr 2*.

- Fernell, E., & Gillberg, C. (2020). Borderline intellectual functioning. In M. J. Aminoff & F. Boller (red.), *Handbook of Clinical Neurology* (Vol. 174, pp. 77–81). Elsevier.
- Fohlin, N. (2024). Det som fattas dagens skola är det besjälade. *Vi Lärare*.
<https://www.vilarare.se/specialpedagogik/inlarning/fohlin-det-som-fattas-dagens-skola-ar-det-besjalade/>
- Folkhälsomyndigheten. (2023). *Skolbarns hälsovanor i Sverige 2021/22. Nationella resultat*. Stockholm: Folkhälsomyndigheten.
- Fredriksson, P., Oosterbeek, H., & Öckert, B. (2012). Långsiktiga effekter av mindre klasser. Uppsala: Institutet för arbetsmarknads- och utbildningspolitisk utvärdering (IFAU).
- Gillberg, C. (2010). The ESSENCE in child psychiatry: Early symptomatic syndromes eliciting neurodevelopmental clinical examinations. *Research in Developmental Disabilities*, 31(6), 1543–1551.
- Gillberg, C. (2014). ESSENCE samlar diagnoserna till en helhet. *Läkartidningen*, 111(CU47).
- Högberg, B., Lindgren, J., Johansson, K., Strandh, M., & Petersen, S. (2021). Consequences of school grading systems on adolescent health: evidence from a Swedish school reform. *Journal of Education Policy*, 36, 84–106.
- Ingvar, M. (2009). Hjärnbarkens funktion. I I. Olsson m.fl. (red.), *Hjärnan* (s. 29–46). Stockholm: Karolinska Institutet University Press.
- Institute for Health Metrics and Evaluation. (2024). Global burden of disease (GBD). Institute for Health Metrics and Evaluation. <http://www.healthdata.org/gbd>
- Joelsson, J. (2015). Konsten att räkna utan att räkna – De automatiserade tabellkunskapernas betydelse för de matematiska strängarnas sammanflätning. Högskolan Kristianstad.
- Klapp A. (2008). *Grades and grade assignment: effects of student and school characteristics*. Avhandling. Institutionen för Pedagogik och Didaktik, Göteborgs universitet.
- Klapp, A. (2015). Tidigare betyg kan sänka svaga elever. *Pedagogiska Magasinet*.
<http://www.lararnasnyheter.se/pedagogiska-magasinet/2015/03/05/tidigare-betyg-kan-sanka-svaga-elever>
- Landgren, V., Lindblad, I., Gillberg, C., & Fernell, E. (2024a). Inability to start or complete upper secondary school strongly predicts unemployment and psychosocial and psychiatric adversities – A register-based follow-up study from Southwestern Sweden. *Acta Paediatrica*. Advance online publication. <https://doi.org/10.1111/apa.17000>
- Landgren, V., Svensson, L., Törnåge, C. J., Theodosios, M., Gillberg, C., Johnson, M. K., Knez, R., & Landgren, M. (2024b). Neurodevelopmental problems, general health, and academic achievements in a school-based cohort of 11-year-old Swedish children. *Acta Paediatrica*, 113(5), 506–516.
- Lindblad, I., Westerlund, J., Gillberg, C., & Fernell, E. (2018). Har alla barn i grundskolan förutsättningar att klara nya läroplanens krav? *Läkartidningen*, 115(6), EY3L.
- Lundahl, C., Hultén, M., Klapp, A., & Mickwitz, L. (2015). Betygens geografi. Delrapport från SKOLFORSK-projektet. Forskning om betyg och summativa bedömningar i Sverige och internationellt. Vetenskapsrådets rapporter. Stockholm: Vetenskapsrådet.
- Miniscalco, C., Fernell, E., Thompson, L., Sandberg, E., Kadesjö, B., & Gillberg, C. (2018). Development problems were common five years after positive screening for language disorders and/or autism at 2.5 years of age. *Acta Paediatrica*, 107(10), 1739–1749.
- Påhlman, M., Gillberg, C., & Himmelmann, K. (2021). Autism and attention-deficit/hyperactivity disorder in children with cerebral palsy: High prevalence rates in a population-based study. *Developmental Medicine & Child Neurology*, 63(3), 320–327.

- Skogstad, I. (2019). *Obekväma sanningar om skolan*. Stockholm: Natur & Kultur.
- Folkhälsomyndigheten (2023). Skolbarns hälsovanor i Sverige 2021/22. Stockholm: Folkhälsomyndigheten.
- Skolverket (2019). Hälsa för lärande – lärande för hälsa, 2019.
<https://www.skolverket.se/publikationsserier/forskning-for-skolan/2019/halsa-for-larande-larande-for-halsa>
- Skolverket (2020). Hur påverkas elever av tidiga betyg och nationella prov?
<https://www.skolverket.se/skolutveckling/forskning-och-utvarderingar/artiklar-om-forskning/hur-paverkas-elever-av-tidiga-betyg-och-nationella-prov>
- Skolverket (2022). Läroplan för grundskolan, förskoleklassen och fritidshemmet Lgr22.
<https://www.skolverket.se/download/18.2cb9d0c18340c7dae31dd/1663664935672/pdf9718.pdf>
- Skolverket (2023a). Elever i Specialskolan läsåret 2022/23.
https://www.skolverket.se/download/18.77dcd21f187010822346cb/1679914750976/pdf113_10.pdf
- Skolverket (2023b). Behörigheten till gymnasieskolan stabil sedan förra året.
<https://www.skolverket.se/om-oss/press/pressmeddelanden/pressmeddelanden/2023-09-28-behorigheten-till-gymnasieskolan-stabil-sedan-forra-aret>
- Skolverket (2024a). Mottagande i anpassade grundskolan, 2024.
<https://www.skolverket.se/regler-och-ansvar/ansvar-i-skolfragor/mottagande-i-anpassade-grundskolan>
- Skolverket (2024b). Utbildning för rörelsehindrade ungdomar (Rh-anpassad utbildning).
<https://www.skolverket.se/regler-och-ansvar/ansvar-i-skolfragor/utbildning-for-rorelsehindrade-ungdomar-rh--anpassad-utbildning>
- Stübner C., Flynn T., Gillberg C., Fernell E., & Miniscalco, C. (2020). Schoolchildren with unilateral or mild to moderate bilateral sensorineural hearing loss should be screened for neurodevelopmental problems. *Acta Paediatrica*, 109(7), 1430–1438.
- Svenska läkaresällskapet (2021). *Kraftsamling för ungas psykiska hälsa*. Stockholm: Svenska läkaresällskapet arbetsgrupp.
- Vygotsky, L.S. (1978). *Mind in Society: The Development of Higher Psychological Processes*. Cambridge, MA: Harvard University Press.
- Whitehurst, G. J. (2003). Presentation by Grover J. Whitehurst, Director, Institute of Education Sciences, IES. Presentation at the Secretary's Mathematics Summit. Washington DC: United States Department of Education.
- Ågren G., & Bremberg, S. (2017). Mortality trends for young adults in Sweden in the years 2000–2017. *Scandinavian Journal of Public Health*, 50(4), 448–453.



Institutionen för pedagogik och specialpedagogik

Konsekvenser av låga och underkända betyg för elevers motivation, prestationer och välbefinnande

Alli Klapp, Göteborgs universitet

Anders Jönsson, Högskolan Kristianstad

David Clarke, Göteborgs universitet

Alli Klapp är docent i pedagogik vid Göteborgs universitet. Hon bedriver forskning om bedömning, betygssättning, och om betydelsen av individuella skillnader hos elever för att lära och prestera. Alli leder infrastrukturen Utvärdering genom uppföljning (UGU) som är en longitudinell enkätstudie. Alli är prefekt för institutionen för pedagogik och specialpedagogik vid Göteborgs universitet.

Anders Jönsson är professor i didaktik vid Högskolan Kristianstad. Han bedriver forskning om bedömning och betygssättning. Anders är för närvarande studierektor för högskolans forskarutbildning i Pedagogiskt arbete.

David Clarke är doktorand i pedagogik vid Göteborgs universitet. Han forskar om bedömning och hans avhandlingsprojekt handlar om konsekvenser av betyg för elevernas lärande och prestationer.

Innehållsförteckning

Förord.....	3
Sammanfattning.....	4
Inledning	5
Kapitel 1	6
Betyg påverkar elevers självbild, engagemang och studiestrategier	6
Betygsomdömen avleder elevernas uppmärksamhet från deras lärande.....	7
Kontinuerlig bedömning upplevs som övervakning och leder till konformitet.....	8
Kapitel 2	10
En absolut gräns för godkända prestationer.....	10
Pedagogiska konsekvenser av en absolut gräns.....	10
Kapitel 3	12
Meritpoäng i det svenska betygssystemet.....	12
Andel elever med underkända betyg uppdelat på olika elevgrupper.....	12
Trendanalyser av underkänt betyg i årskurs 6 och hur det går det i årskurs 9 och 12.....	14
Förskjuten betygsskala – ekvidistant och kontinuerlig betygsskala	18
Kapitel 4	22
4.1 Oavsiktligt utfall av underkända betyg	23
4.2 Vad i stället?	23
Slutord.....	24
Referenser	24–26

Förord

Hur påverkar betyg elevers lärande? Kan betygssystemet stärka att likvärdiga betyg sätts på elevers kunskaper? Vilka möjligheter har elever med underkända betyg i grundskolan att studera vidare på gymnasiet? Dessa frågor om betyg, betygssystemets konstruktion och underkända betyg bör fördjupas och diskuteras för att förstå hur betygssystemet påverkar unga människor möjlighet till utbildning och arbete.

Betyg och bedömning är ett område som är belastat med många föreställningar och där antaganden tas för givna. Det finns därför ett behov av kunskap om konsekvenser av underkända betyg för elevernas lärande, prestationer, motivation och välmående och för betyg som urvalsinstrument. I denna rapport diskuteras forskning om hur låga och underkända betyg påverkar elevers prestationer, motivation och välmående. Deskriptiva analyser av två stora datamaterial presenteras och ger en grund för att diskutera konsekvenser av underkända betyg för elever och för utbildningssystemet i stort.

Sammanfattning

Denna rapport diskuterar godkäntgränsen i den svenska betygsskalan dels med utgångspunkt från tidigare nationell och internationell forskning om pedagogiska konsekvenser av låga och underkända betyg, dels utifrån registerdata och trendanalyser samt longitudinella analyser av elever med låga och underkända betyg och deras väg genom utbildningssystemet.

Resultaten från tidigare forskning om pedagogiska konsekvenser av låga och underkända betyg visar att elever som får låga betyg tenderar att tappa tilltron till den egna förmågan, vilket kan leda till minskad motivation och ökad användning av mindre konstruktiva studiestrategier, vilket i sin tur påverkar skolprestationerna negativt. Detta kan leda till en negativ spiral, där ytterligare försämringar i prestationer ger ännu lägre betyg, om inte eleven bl.a. får det stöd som krävs för att vända utvecklingen. Ett antal svenska studier tyder även på att tidigare betygssättning och ett ökat fokus på betyg kan leda till minskad tillfredsställelse med livet och ökad sannolikhet för diagnostiserad ohälsa, framför allt i form av depression och ångest hos flickor. Tidigare forskning visar också att betyg kan avleda elevernas uppmärksamhet från lärande och ämneskunskaper, mot social jämförelse och sociala strategier, som att hela tiden ”visa upp sig från sin bästa sida” gentemot betygssättande lärare. Flera av dessa konsekvenser kan tänkas bli förstärkta av en absolut gräns för godkända prestationer, eftersom underkända betyg utgör ett hot mot elevernas fortsatta utbildning, men också ett starkare socialt stigma, jämfört med låga betyg. Även om risken för underkända betyg skapar incitament för eleverna att anstränga sig, kan ökad ansträngning inte i sig självt förbättra elevernas prestationer i avsaknad av adekvat stöd. Underkända betyg måste därför, för att ge effekt, kombineras med uppföljning och adekvat stöd.

Svenska data som presenteras i rapporten visar emellertid att antalet elever som får underkända betyg i årskurs 6 är i stort sett detsamma som i årskurs 9, vilket tyder på att det stöd som ges inte är tillräckligt och/eller adekvat. En del av de elever som fått betyget F i årskurs 6 får ett godkänt betyg (E) i årskurs 9, men en annan grupp elever, som fått godkända betyg i årskurs 6, får underkänt betyg i årskurs 9. Det finns även en grupp elever som fått höga betyg i årskurs 6, men som ändå fått betyget F i årskurs 9.

Inledning

Denna rapport behandlar båda svensk och internationell forskning om låga och underkända betyg och dess konsekvenser på elevernas motivation, prestationer och välbefinnande. Det är inte en systematisk översikt som följer en viss metod eller design för att försöka hitta all forskning om underkända betyg. Vi har gjort sökningar i databaser och sökt i nationella och internationella forskningstidskrifter som behandlar bedömningsfrågor. Utifrån vår egen expertis inom området har vi kunskap om forskare och forskargrupper som bedriver forskning inom bedömningsområdet, resultat som vi också inkluderat i denna rapport.

Rapporten är indelad i fyra delar. Eftersom forskningen på området primärt intresserat sig för konsekvenser av låga betyg, ges först en allmän översikt av forskningsläget kring hur låga betyg tenderar att påverka elevers motivation, prestationer och välbefinnande (kapitel 1). Dessa resultat diskuteras därefter i relation till underkända betyg, med fokus på vilka konsekvenser som kan tänkas bli förstärkta eller mildras av en absolut gräns för godkända prestationer (kapitel 2). I kapitel 3 presenteras sammanställningar av empiriska data från Sverige, som hämtats från projektet Utvärdering genom uppföljning (UGU) som ansvarar för en omfattande longitudinell databas vid Göteborgs universitet samt Gothenburg Educational Longitudinal Database (GOLD). Slutligen, i kapitel 4, presenteras slutsatser från tidigare kapitel.

Kapitel 1: Forskningsläget: Hur påverkar låga betyg elevers motivation, prestationer och välbefinnande

Det kan vara svårt att skapa sig en tydlig bild av den internationella forskningen kring pedagogiska konsekvenser av betyg, vilket bl.a. beror på gränsdragningen mellan olika typer av återkoppling. Om man ska vara strikt, så är betyg (i Sverige) ett standardiserat omdöme som ges exempelvis i slutet på grundskolan eller i slutet på en kurs i gymnasieskolan. Många gånger ges emellertid betygslänkande omdömen även som prognoser under terminens eller kursens gång, och ofta även som återkoppling på enskilda uppgifter. Även om betygslänkande omdömen formellt sett inte är betyg, kan de ge motsvarande pedagogiska effekter. Eftersom de ges oftare, förekommer de även mer frekvent i forskningen.

Summativa omdömen, som betyg och betygslänkande omdömen, kan även ges utifrån olika skalor och utifrån olika bedömningsprinciper. I Sverige har vi för närvarande en skala från F till A, men vi har historiskt haft flera olika betygsskalor. Andra länder har skalor som, inte sällan av historiska skäl, ser annorlunda ut. Exempelvis i Danmark går den nuvarande skalan från -3 till +12, dock inte kontinuerligt, utan i sju steg. I dagsläget har flera länder, inklusive Sverige, någon form av kriterierelaterad betygssättning, men det finns även exempel på att normrelaterad bedömning används för standardiserade prov som t.ex. högskoleprovet. Sverige hade också normrelaterade betyg under stora delar av 1900-talet.

Den stora spridningen mellan olika skalor och bedömningsprinciper gör det viktigt att diskutera i vilken utsträckning det är möjligt att generalisera resultat från studier genomföra i andra kontexter. Skulle man enbart titta på studier som undersökt precis samma typ av betygssystem som det nuvarande svenska, skulle underlaget vara skralt och knappast lämpa sig för att dra några tydliga slutsatser. Vid sammanställningar av studier från olika kontexter kan man emellertid se stora likheter i de pedagogiska konsekvenserna, trots spridningen mellan olika

skalor och bedömningsprinciper. Något som tycks avgörande är att återkopplingen ges utifrån någon form av skala, vilket – oavsett skalans längd eller skalsteg – ger möjlighet till rangordning och jämförelser elever emellan. Detta skiljer sig markant från kvalitativ återkoppling, som beskriver styrkor och utvecklingsmöjligheter i elevernas prestationer, men utan att placera dessa längs en skala. En annan faktor som påverkar betygens konsekvenser, är hur viktiga betygen är för eleverna. I Sverige är betygen s.k. *high-stakes*, eftersom de kan ha avgörande betydelse för elevernas framtid, framför allt genom att öppna eller begränsa elevernas möjligheter till fortsatt utbildning.

Det man därför bör ha med sig som läsare av denna rapport är att underlaget visserligen bygger på ett heterogent underlag vad gäller olika skalor och bedömningsprinciper (även om särskild vikt har lagts vid att lyfta fram aktuella svenska studier), vilket gör att man får vara försiktig med generaliseringar av enskildheter, men att resultaten kring pedagogiska konsekvenser trots detta är relativt entydiga, eftersom alla former av betyg och betygsliknande omdömen ges i förhållande till en skala.

Betyg påverkar elevers självbild, engagemang och studiestrategier

Kiefers studie från 1975 har i flera sammanhang lyfts som ett illustrativt exempel på hur elever kan reagera på att få betyg (t.ex. Skolverket, 2024). I denna studie lät man elever i flera olika årskurser (2, 4, 6 och 8) fylla i enkäter, som handlade om deras självförtroende och deras tilltro till den egna förmågan. Man valde dessutom att specifikt fråga elever som var antingen hög- eller lågpresterande i respektive årskurs. Genom denna design kunde man undersöka skillnader mellan hög- och lågpresterande elever i hur de reagerade på att få betyg, från tidig skolålder till tonåren. Resultaten visar bl.a. att det i årskurs 2 inte fanns någon större skillnad mellan låg- och högpresterande elever när det gällde tilltron till den egna förmågan, men att detta ändrades över tid. De högpresterande elevernas tilltro till den egna förmågan låg kvar på samma höga nivå i högre årskurser, medan den sjönk stadigt för de lågpresterande eleverna. Kurvan för det allmänna självförtroendet gick inte lika tydligt nedåt för lågpresterande elever, men låg konstant långt under nivån för högpresterande elever. Det fanns även skillnader mellan pojkar och flickor, där lågpresterande flickor genomgående skattade sitt självförtroende och tilltro till den egna förmågan lägre än pojkarna, men dessa skillnader var över lag betydligt mindre jämfört med skillnaden mellan hög- och lågpresterande elever (oavsett biologiskt kön).

Liknande studier har även genomförts i Sverige. Klapp (2015) jämförde t.ex. hur det gick för elever som fått – respektive inte fått – betyg i årskurs 6. Studien hämtade data från databasen "Utvärdering genom uppföljning" (UGU), som innehåller enkät- och registerdata för ett nationellt representativt urval av individer födda mellan 1948 och 1998. I analysen ingick data från 8 558 elever från 430 klasser och 29 skolhuvudmän och även här visade resultaten att betygssättning i årskurs 6 generellt hade en negativ inverkan på elevernas prestationer på högstadiet och i gymnasieskolan. Framför allt var det de lågpresterande eleverna bland dem som fick betyg i årskurs 6 som drabbades, genom att de fick en sämre utveckling på högstadiet, och hade lägre sannolikhet att slutföra sina gymnasiestudier, jämfört med motsvarande elevgrupp som inte fått betyg i årskurs 6. Det fanns en tendens att högpresterande elever, främst flickor, gynnades av att betygssättas, men denna effekt var svag.

Det som dessa studier visar är att låga betyg påverkar elevernas självkänsla och tilltro till den egna förmågan negativt, vilket leder till en sämre utveckling för dessa elever över tid. En viktig anledning till denna utveckling är att elevernas självkänsla och tilltro till den egna förmågan är sammankopplad med deras engagemang och studiestrategier. Om eleverna inte tror att de har

förmågan att klara uppgifterna i ett visst ämne (eller i skolan mer allmänt), ökar risken för att de använder sig av mindre konstruktiva studiestrategier och minskar sitt engagemang i skolarbetet. Det finns ingen större anledning att anstränga sig om man ändå inte har möjlighet att lyckas.

Ett typiskt exempel på mindre konstruktiva studiestrategier är s.k. ”undvikandestrategier”. Dessa strategier handlar om att eleven, i stället för att lägga fokus på att lära sig, försöker hitta sätt för att undvika att misslyckas eller visa sin okunskap. Det kan t.ex. handla om att inte engagera sig i skolarbetet, eller låta bli att slutföra eller lämna in uppgifter. Att eleven undviker att visa sina brister, innebär även att läraren inte uppmärksammas på elevens svårigheter och därför inte kan ge adekvat stöd. Undvikandestrategier leder därmed till att elevens möjligheter att lära sig minskar, vilket i sin tur ökar risken för låga eller underkända betyg ännu mer, och ytterligare försämring av tilltron till den egna förmågan. Ett antal storskaliga svenska studier tyder även på att tidigare betygssättning och ett ökat fokus på betyg kan leda till minskad tillfredsställelse med livet och ökad sannolikhet för diagnosticerad ohälsa, framför allt i form av depression och ångest hos flickor (Collins & Lundstedt, 2021; Högberg m.fl., 2021; Linder m.fl., 2023).

Betygsomdömen avleder elevernas uppmärksamhet från deras lärande

Det finns en utbredd föreställning om att betyg och betygsliknande omdömen skapar motivation hos eleverna och att sådana omdömen kan hjälpa eleverna att förbättra sina prestationer i skolan (Harlen & Deakin-Crick, 2002). Den empiriska forskningen ger emellertid inte riktigt samma bild. I en meta-analys, där man har sammanställt resultat från flera studier på detta tema, framkommer att betygsliknande omdömen visserligen kan ge visst stöd för att eleverna ska kunna förbättra sina prestationer, jämfört med att inte få någon återkoppling alls. Betygsliknande omdömen ger dock betydligt sämre stöd för såväl elevers motivation som för deras prestationer, jämfört med mer detaljerad återkoppling. Elever som får formativ återkoppling har högre motivation och mer intresse, jämfört med elever som får betyg, och de upplever sig dessutom i högre grad som kompetenta och presterar bättre (Koenka m.fl., 2021). Betyg och betygsliknande omdömen tycks således snarare vara ett hinder för elevernas motivation och kunskapsutveckling än ett stöd.

Man skulle kunna tänka sig att läraren borde ge både formativ återkoppling och betygsliknande omdömen, för att komma runt problemet med de negativa effekterna av betyg och betygsliknande omdömen. Tyvärr verkar det inte vara så enkelt, utan forskning på området tyder i stället på att elever i sådana situationer tenderar att främst fokusera på betygsomdömet, inte på den formativa återkopplingen. Det förmodligen mest kända exemplet är från en studie av Butler (1988), som jämförde effekterna av tre olika typer av återkoppling för elever i mellanstadiet (11 år): (1) kommentar på det eleverna gjort, följt av förslag på förbättring; (2) enbart betygsomdöme; (3) både kommentar och betygsomdöme. I studien klarade sig de elever som fick kommentarer på vad de gjort, men inget betygsomdöme, bättre än de andra, och de behöll också sitt intresse för materialet de arbetat med (eller så ökade intresset). Sämst gick det för de elever som fick både kommentar och betygsomdöme. De elever som fick låga betygsomdömen tenderade att tappa intresset, och kommentarerna kunde inte påverka detta. Intressant att notera är dessutom att över hälften av de elever som fick både kommentar och betygsomdöme inte kunde komma ihåg kommentarerna efteråt, jämfört med ungefär 90 procent av de som enbart fick kommentarer. Detta tyder på att de elever som fick både kommentar och betygsomdöme främst fokuserade på betygsomdömet, inte på kommentarerna.

Det finns olika förslag på förklaringar till att elever reagerar på detta sätt. Ett betygsomdöme upplevs t.ex. inte enbart som ett kvitto på hur väl man har presterat, utan oftast mer personligt än så. Betygsomdömen riktar därmed elevernas uppmärksamhet mot dem själva (s.k. ”*ego-involvement*”) i stället för mot uppgiften (”*task-involvement*”). Om dessa känslor dessutom påverkar självbilden negativt, så att tilltron till den egna förmågan minskar, riskerar man – precis som beskrevs i föregående avsnitt – att eleven minskar sitt engagemang i skolarbetet (Black m.fl., 2004). En annan förklaring handlar om att (som beskrivs ovan) omdömen som kan placeras längs en skala underlättar jämförelser mellan elever, och att sådana omdömen därmed tenderar att flytta elevernas uppmärksamhet från ämnesinnehållet till social jämförelse (dvs. om man presterade bättre eller sämre än andra). Kvalitativ återkoppling, i form av styrkor och utvecklingsmöjligheter, kan inte jämföras på samma sätt och kan därför få eleverna att i högre grad fokusera på ämnesinnehållet.

Kontinuerlig bedömning upplevs som övervakning och leder till konformitet

Skolreformerna från 2011 och framåt har skapat ett kraftigt ökat fokus på bedömningar för summativa syften i det svenska skolsystemet (se t.ex. Löfgren m.fl., 2021; Skolverket, 2015; Wahlström & Sundberg, 2015). Som påpekats tidigare, finns det även studier som tyder på att införandet av betyg från en tidigare ålder är förknippat med ökad skolrelaterad stress och minskad akademisk självkänsla bland elever, vilket lett till en ökning av psykosomatiska symtom och minskad tillfredsställelse med livet (Högberg m.fl., 2021; Linder m.fl., 2023).

Intervjustudier med elever i olika stadier av skolsystemet (t.ex. Pérez Prieto & Löfgren, 2017; Hirsh, 2020; Löfgren m.fl., 2021) ger också en relativt samstämmig bild av hur bedömningen är ständigt närvarande i medvetandet hos svenska elever, både i och utanför klassrummet. Elever påstår också att de känner sig ständigt övervakade av sina lärare, vilket lämnar lite utrymme för att ställa frågor eller göra misstag, utan att samtidigt oroa sig för hur det kan påverka betygen (t.ex. Rönn & Pettersson, 2023).

För att hantera denna situation, där eleverna ständigt bedöms, försöker eleverna att förbättra eller säkra sina betyg genom att förändra sitt beteende och påverka sina sociala relationer med lärarna. Även om det förmodligen skulle vara mer effektivt för eleverna att lägga sina resurser på studierna, verkar de inte tycka att förväntningarna är tillräckligt tydliga för att uteslutande förlita sig på sina akademiska prestationer. Följaktligen investerar de även i sociala relationer, som en viktig del av sina studiestrategier. Exempelvis beskriver elever i en studie av Löfgren och Löfgren (2016), som utvärderat elevers erfarenheter av att få betyg i årskurs 6, vikten av att vara uppmärksam på vad läraren vill för att få högre betyg, som att vara väluppfostrad och göra det man blir tillsagd. På liknande sätt hävdar elever i årskurs 9 att deras betyg inte uteslutande återspeglar kvaliteten på deras prestationer, utan också används som belöning för ansträngning och önskvärdt beteende (Vogt, 2017). En sådan bedömningskultur kan skapa stark stress hos eleverna och i utvärderingen av försöksverksamheten med betygssättning i årskurs 4 (Löfgren m.fl., 2021) vittnar många elever om stressen av att ha arbetat sig upp till ett högt betyg, men samtidigt vara konstant oroad över att minsta felsteg kan sänka betyget igen.

Kapitel 2: Konsekvenser av en skarp godkäntgräns

De forskningsresultat som sammanfattas i kapitel 1 kommer från en bredd av olika betygssystem, även om särskild vikt läggs på aktuella svenska studier. Det är dock inte alltid tydligt om det finns en absolut gräns för godkända prestationer i respektive system, eller vilken

roll denna gräns i så fall spelar. I det här kapitlet diskuterar vi därför resultaten i relation till underkända betyg, med fokus på vilka konsekvenser som kan tänkas bli förstärkta eller mildras av en absolut gräns för godkända prestationer. Först görs några förtydliganden kring innebörden av en godkänthgräns.

En absolut gräns för godkända prestationer

Betyget för godkänt resultat (E) i det svenska betygssystemet är "absolut", vilket innebär att *samtliga kriterier* måste vara uppfyllda för att eleven ska uppnå betyget E. Elever som misslyckas med att uppnå något av de angivna kriterierna blir därmed underkända och får F i betyg. Denna typ av absolut bedömning kallas för "analytisk disjunktiv betygsregel" och innebär att läraren inte kan låta elevens styrkor kompensera för svagare sidor vid betygssättningen (Sadler, 2005, 2009). Denna regel används inom vissa yrkesutbildningar och några ämnesdiscipliner inom högre utbildning. Ett tydligt exempel på när denna absoluta regel är rimlig, är när studenter inom medicin bedöms på moment där de måste behärska en viss teknik för att operera patienter. Dessa studenter kan inte kompensera för brister i förmågan att operera med t.ex. förmågan att kommunicera med patienten. Det finns en absolut gräns för kunskaper som läkare måste behärska och detta kan inte kompenseras med andra kunskaper eller kompetenser. När det handlar om liv och död är en disjunktiv betygsregel således önskvärd, men den är inte självklar i andra sammanhang.

Utgångspunkten att alla kriterier ska vara uppfyllda för ett godkänt betyg bygger på antaganden som inte nödvändigtvis stämmer. En central utgångspunkt är t.ex. att det – med rimlig precision – går att avgöra vilka elever som hamnar på respektive sida om godkänthgränsen. Tillförlitligheten i denna typ av gränsdragningar är mycket väl undersökta inom pedagogisk forskning. Forskningen visar att sådana gränser har en hög grad av osäkerhet, även i förhållande till standardiserade, skriftliga prov (se t.ex. Koretz, 2008). I utredningen *Likvärdigt, rättssäkert och effektivt – ett nytt nationellt system för kunskapsbedömning* (SOU 2016:25, Del 2, s. 162) redovisas resultat från en utvärdering av osäkerheten i de svenska nationella proven i matematik, där man uppskattar att 20–30 procent av eleverna får ett för högt eller för lågt provbetyg i förhållande till deras faktiska kunskaper. Med tanke på att betygssättning dels innebär ett större och mer heterogent underlag jämfört med provresultat, dels att betygsriterierna är betydligt mer öppna för tolkning än bedömningsanvisningar till nationella prov, är siffran troligen avsevärt högre för betyg än för provbetyg. Experimentella studier som undersökt överensstämmelsen i lärares betygssättning, visar på en överensstämmelse på 45–65 procent (Jönsson & Balan, 2018; Jönsson m.fl., 2020). Uppemot varannan elev skulle därmed kunna få ett annat betyg, om det var en annan lärare som satte betygen. Det är således uppenbart att vi egentligen inte vet vilka elever som faktiskt uppnår godkända prestationer.

Pedagogiska konsekvenser av en absolut gräns

Som framgår av föregående kapitel, tyder forskning kring betyg och betygslänkande omdömen på att betyg kan ha en rad pedagogiska konsekvenser. Centralt är t.ex. att elever som får låga betyg tenderar att tappa tilltron till den egna förmågan, vilket kan leda till minskad motivation och ökad användning av mindre konstruktiva studiestrategier; detta i sin tur påverkar skolprestationerna negativt. Ett antal svenska studier tyder även på att tidigare betygssättning och ett ökat fokus på betyg kan leda till minskad tillfredsställelse med livet och ökad sannolikhet för diagnostiserad ohälsa, framför allt i form av depression och ångest hos flickor (Collins & Lundstedt, 2021; Högberg & Strandh, 2024 Linder m.fl., 2023). Frågan i det här avsnittet är hur en absolut gräns för godkända prestationer kan tänkas påverka betygens pedagogiska konsekvenser.

Ett antagande, som ligger bakom utgångspunkten om att alla kriterier ska vara uppfyllda för ett godkänt betyg, är att ett krav på en miniminivå av kunskaper leder till att elever lär sig mer (Levin, 2007). Man kan t.ex. tänka sig att skolan (dvs. lärare, skollledning och/eller elevhälsa) reagerar på att vissa elever riskerar att inte uppnå godkända resultat, och därmed sätter in de resurser som krävs för att stötta dessa elever, så att de kan nå åtminstone miniminivån. I utvärderingen av försöksverksamheten med betygssättning i årskurs 4 ges ett antal exempel på detta (Löfgren m.fl., 2021). Urvalet av skolor i denna utvärdering är dock knappast representativt för landet som helhet, och den generella bilden skiljer sig påtagligt genom att en mycket stor andel av de som får betyget F på mellanstadiet är fortsatt underkända i årskurs 9, vilket vi kommer att återkomma till i nästa kapitel.

Man kan också tänka sig att ett krav på en miniminivå av kunskaper leder till att elever lär sig mer, eftersom de negativa effekterna av att misslyckas blir så mycket större jämfört med att få ett lågt betyg. I dagens system kan elever t.ex. bli helt utestängda från vidare utbildning om de underkänns i vissa ämnen, till skillnad från system där antagningen baseras på ett betygsmedelvärde och där låga betyg i vissa ämnen således kan kompenseras av höga betyg i andra. Det är möjligt att hotet om underkända betyg kan ge starkare incitament för eleverna att anstränga sig i skolan, även om det inte finns några tydliga vetenskapliga belägg för detta antagande. Det största problemet ligger emellertid i antagandet att dessa incitament, på egen hand, skulle bidra till förbättrade prestationer. En central anledning till att många elever inte förmår uppnå godkända resultat är att de saknar de resurser som krävs, t.ex. i form av stöd (från skolan eller hemmet), men också i form av mindre påtagliga resurser som självförtroende, tidigare kunskaper och kognitiva förmågor. I utvärderingen av försöksverksamheten med betygssättning i årskurs 4 (Löfgren m.fl., 2021), ges ett antal exempel på hur högpresterande elever har ett batteri av tillgängliga resurser att tillgripa vid behov, medan lågpresterande elever uppger att de behöver "arbeta hårdare", men utan att kunna ge några konkreta exempel på hur de kan göra (Jönsson, 2022).

Som beskrevs i föregående kapitel, är elevernas självkänsla en grundläggande anledning till flera av de negativa konsekvenser som dokumenterats som en följd av låga betyg. Även här kan man tänka sig att effekterna förstärks av att eleverna får underkända betyg, snarare än att få låga betyg, genom att underkända betyg väcker starkare negativa känslor hos eleverna (s.k. *"achievement emotions"*; se Pekrun m.fl., 2007). Starkare negativa känslor kan i sin tur leda till en ökad sannolikhet för eleverna att ge upp eller på annat sätt avleda deras uppmärksamhet från lärande och ämneskunskaper mot social jämförelse eller sociala strategier.

Kapitel 3: Den empiriska bilden

Utifrån elevdata i databaserna Gothenburg Educational Longitudinal Database (GOLD) och inom projektet Utvärdering genom uppföljning (UGU) presenteras i detta avsnitt resultat för hur det går för elever med låga och underkända betyg i grundskolan och i gymnasieskolan. GOLD är en databas som innehåller populationsdata, vilket innebär att alla individer finns med för respektive år och årskull, medan UGU innehåller enkätdata för ett urval av elever i Sverige från 1961 och framåt. I analyserna som presenteras nedan ingår data som hämtats från GOLD för totalt 1 036 213 elever (51,5 % pojkar och 48,5 % flickor), som gick i skolan 2013–2020, samt data från UGU för cirka 10 000 elever födda 2004, vilket utgör ett riksrepresentativt urval.

Meritpoäng i det svenska betygssystemet

När elever använder sina betyg för att söka vidare inom utbildningssystemet, översätts betygen till siffervärden. Nuvarande poängsystem innebär att godkänt betyg (E) ger 10 poäng, varefter varje steg längs skalan ger ytterligare 2,5 poäng. Betyget D ger således 12,5 poäng, C 15 poäng, B 17,5 poäng och A 20 poäng. Betyget F avviker emellertid från ovanstående mönster genom att ge 0 poäng, vilket ger ett särskilt stort avstånd mellan just underkänt och godkänt betyg, jämfört med övriga skalan. Eftersom gränsen för betyget E (till skillnad från övriga betygssteg) dessutom är absolut, är skalan varken ekvidistant eller kontinuerlig. Att få betyget F innebär således att eleven kan missgynnas både genom att utestängas från vidare studier, eftersom det krävs godkända betyg i vissa ämnen, och genom att meritpoängen påverkas särskilt kraftfullt på grund av det stora avståndet i poäng mellan betygen F och E.

Förutom problemen kopplade till svårigheterna att komma vidare i utbildningssystemet, finns det även en risk för oönskade konsekvenser av att betyget F i ett ämne ger 0 poäng, eftersom det signalerar att eleven inte har några kunskaper i ämnet, även om eleven har deltagit i undervisningen. Att elevens kunskaper inte når upp till betygskriterierna för betyget E, innebär emellertid inte att eleven helt saknar kunskaper i ämnet.

Andel elever med underkända betyg uppdelat på olika elevgrupper

I detta avsnitt presenteras data från populationen elever som gick i skolan i årskurserna 6, 9 och 12 (dvs. år 3 i gymnasieskolan) perioden 2013–2020. Beroende på vilken årkurs som avses kan data för vissa år saknas, t.ex. på grund av att elever inte befunnits sig i en viss årskurs.

Tabell 1 visar andelar av elever inom olika elevgrupper som fått minst ett underkänt betyg (F) i årskurs 6, 9 eller 12. I tabellens översta del är det den totala andelen av elever som fått minst ett underkänt betyg någon gång under sin skoltid och därefter redovisas andelar per år. Könsskillnader redovisas därunder och dessa är relativt stabila över perioden. Nästa del visar på avsevärda skillnader mellan utländsk och svensk bakgrund och de sista raderna i tabellen visar på skillnader med avseende på föräldrars utbildningsnivå. Resultatet visar på den ökande andelen elever som fått minst ett underkänt betyg i gymnasieskolan, jämfört med hur det ser ut i grundskolan. Sammanfattningsvis ser vi att det är pojkar, elever med föräldrar med låg utbildningsnivå samt elever som har utländsk bakgrund som fått minst ett underkänt betyg under sin skoltid, oavsett skolförhållanden.

Tabell 1. Minst ett underkänt betyg för olika elevgrupper 2013–2020. Andelar (%).

	Åk 6	Åk 9	Åk 12
Fått minst ett F totalt 2013–2020	19	21	37
2013	17	20	*
2014	19	20	35
2015	18	20	37
2016	18	20	37
2017	20	21	37
2018	*	21	37
2019	*	22	37
2020	*	22	37
Fått F, flickor	16	18	32
Fått F, pojkar	22	23	42
Fått F, svensk bakgrund	13	16	36
Fått F, utländsk bakgrund	40	47	51
Fått F, ingen högre utbildning	27	29	43
Fått F, >2 år högre utbildning	10	10	30

Källa: GOLD.

Tabell 2 visar hur elever med utländsk bakgrund, uppdelat utifrån vilken ålder eleven kom till Sverige och började skolan. Första gruppen är elever som är födda i Sverige med två utlandsfödda föräldrar. Därefter visas elever som är födda utomlands och som kom till Sverige när de var mellan 0–6 år, de som kom hit när de var mellan 7–12 år och de som kom vid 13–16 års ålder. Sammanfattningsvis visar tabellen att de barn som föddes i Sverige, samt kom till Sverige i åldern 0–6 år, får färre underkända betyg under skoltiden. Elever som kom till Sverige senare, mellan 7–12 och 13–16 års ålder, har i större utsträckning fler än tre underkända betyg.

Tabell 2. Andel elever med minst tre underkända betyg med olika nivåer av utländsk bakgrund. Mellan 2013–2020. Andel (%).

	Åk 6	Åk 9	Åk 12
Född i Sverige med två utländska föräldrar	13	13	25
Flyttade till Sverige 0–6 år	14	14	26
Flyttade till Sverige 7–12 år	41	27	28
Flyttade till Sverige 13–16 år	61	57	23

Källa: GOLD.

Tabell 3 visar andelen elever med underkända betyg i kommunala och fristående skolor under perioden 2014–2020. Skillnaderna är stabila och visar på att andelen elever med fler än tre underkända betyg är större i kommunala jämfört med fristående skolor. I gymnasieskolan ändrar sig dock fördelningen och där är andelen elever med minst tre underkända betyg högre för fristående skolor.

Tabell 3. Andel elever med minst tre underkända betyg i kommunal respektive fristående skola. Mellan 2014–2020. Andel (%).

	Åk 6	Åk 9	Åk 12
Kommunal skola, totalt	9	12	14
Fristående skola, totalt	5	5	16
Kommunal 2014	8	10	14
Fristående 2014	5	5	17
Kommunal 2015	8	11	15
Fristående 2015	5	5	15
Kommunal 2016	9	11	15
Fristående 2016	5	5	16
Kommunal 2017	10	12	14
Fristående 2017	5	5	16
Kommunal 2018		12	14
Fristående 2018		5	17
Kommunal 2019		13	14
Fristående 2019		5	15
Kommunal 2020		13	14
Fristående 2020		6	15

Källa: GOLD.

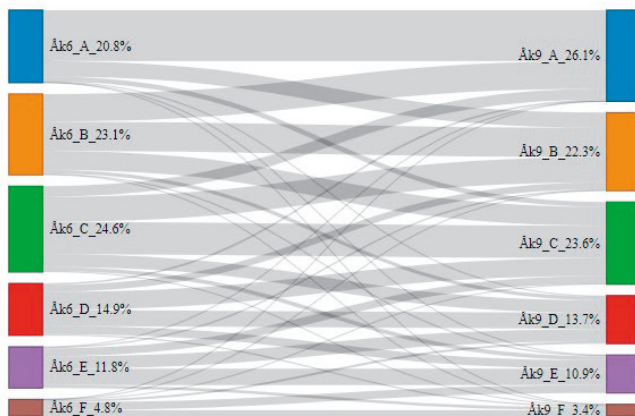
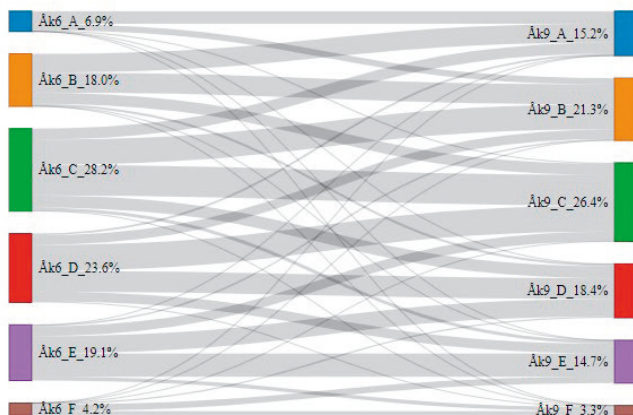
Trendanalyser av underkänt betyg i årskurs 6 och hur det går det i årskurs 9 och 12

I detta avsnitt presenteras resultat för elever som ingår i databasen Utvärdering genom uppföljning (UGU). Elever födda 2004 ingår i urvalet och detta är ett 10-procentigt nationellt representativt urval (ca 10 000 elever). Först presenteras resultat enbart för betyg i årskurs 6 och 9, därefter även bakgrundsfaktorer, som kön och immigration, för att bättre förstå hur olika undergrupper av elever som får underkända betyg i årskurs 6 presterar i årskurs 9.

Figurerna 1–3 visar hur elevernas betyg förändras från årskurs 6 till årskurs 9. Ungefär hälften av eleverna som får underkända betyg i årskurs 6 höjer sina betyg, mestadels till godkänt betyg (E). En del uppnår t.o.m. så höga betyg som B. Det visar sig också att vissa elevers betyg sjunker drastiskt från årskurs 6 till årskurs 9, några få elever faller från betyg A eller B till ett underkänt F-betyg. Detta bekräftar att det är en liten andel elever med F-betyg, i årskurs 6 eller årskurs 9, som är mer kapabla än vad deras betyg anger, eftersom de antingen går vidare till (eller tidigare har uppnått) mycket högre betyg. Ytterligare forskning krävs för att fastställa de bakomliggande orsakerna till varför dessa elever visar så stora betygsvariationer. Det finns många möjliga faktorer, både i och utanför skolan, som kan vara orsak till dessa variationer och som påverkar betygen i antingen årskurs 6 eller i årskurs 9.

Av eleverna som fick F i årskurs 6 i svenska fick drygt 50 procent av eleverna betyget E i årskurs 9, medan 34 procent behöll sitt F och ytterligare 9 procent som tidigare hade högre betyg fick underkänt betyg.

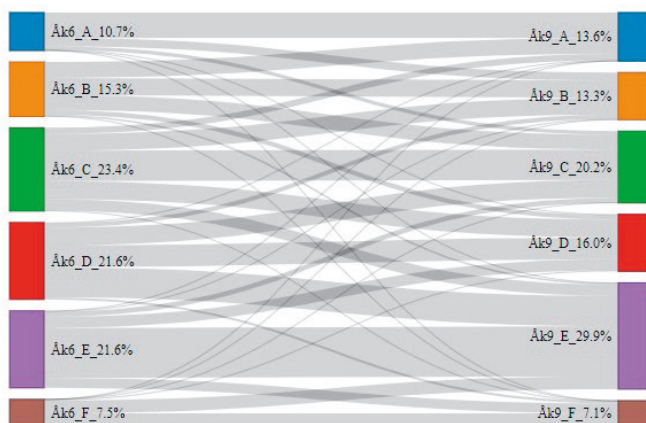
Figur 1. Betygsutveckling från årskurs 6 till 9 i svenska. Siffror anger andel elever för de olika betygsstegen (Källa: UGU).



fick betyget E i årskurs 9. Av eleverna som fick godkända betyg i årskurs 6 fick cirka 11 procent F i årskurs 9.

I samtliga tre ämnen – svenska, engelska och matematik – kommer den största ökningen av nya underkända betyg i årskurs 9 från elever som nådde betyget E i årskurs 6. Att betyget F skulle leda till en ökad motivation hos eleverna är därmed tveksamt, eftersom i stort sett samma antal elever får ett F i årskurs 9. Att ett F antas motivera elever att anstränga sig mer för att lära sig och prestera högre är således inte en mekanism som framträder i elevdata.

Figur 3. Betygsutveckling från årskurs 6 till 9: matematik. Siffror anger andel elever för de olika betygsstegen (Källa: UGU).

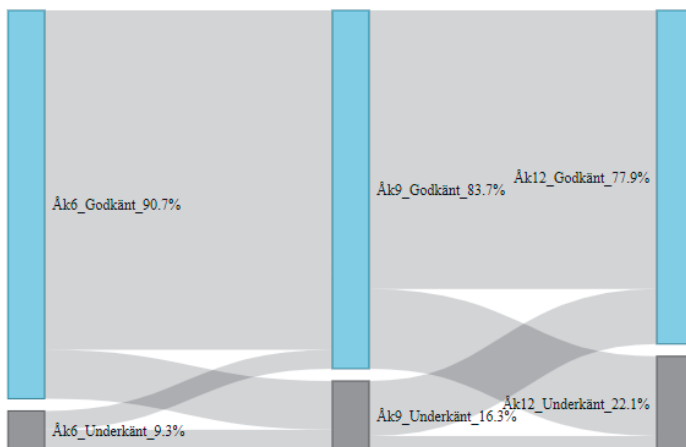


Av de elever som fick underkänt betyg i årskurs 6 i matematik fick 47 procent F i årskurs 9. Cirka 48 procent av eleverna som fick betyget F i årskurs 6 fick betyg E i årskurs 9. Cirka 13 procent av eleverna som fick betyget F i årskurs 9 hade godkända betyg i matematik i årskurs 6.

I samtliga tre redovisade ämnen är mönstret likartat. Tidiga misslyckanden kan i sig vara en bidragande orsak till fortsatt misslyckande (Allodi m.fl., 2010). Elever som får betyget F trots att de anstränger sig, riskerar att uppleva att det inte är lönt att anstränga sig ytterligare för att få ett godkänt betyg, och därmed tappa sin motivation. En viktig bidragande orsak till detta kan vara att betygskriterierna för betyget E är absoluta, dvs. att *alla* kriterier måste vara uppfyllda för att få godkänt betyg. Denna bristande möjlighet för lärare att i betygssättningen kompensera för elevernas svagheter, innebär en uppenbar risk för att elever ger upp om de inte ser det möjligt att uppnå godkända resultat i förhållande till vissa (även enskilda) kriterier.

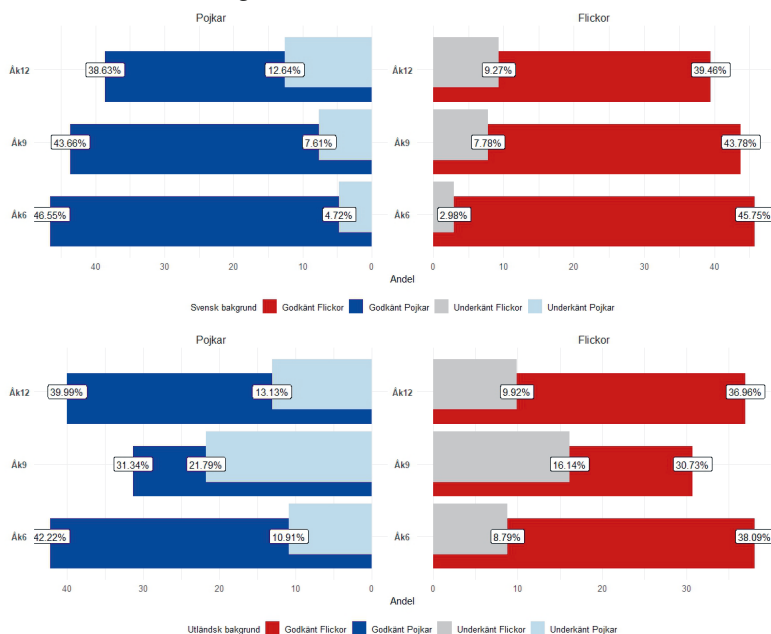
För att bättre förstå variationen i hur betygen fördelas kommer vi i grafer/tabeller att presentera resultat där elevernas bakgrundskaraktäristika såsom kön och immigrationsbakgrund är med i tabellerna.

Figur 4. Andel elever som fick minst ett underkänt betyg i årskurser 6, 9 och 12 (Källa: GOLD).



Figur 4 visar att andelen elever som fått minst ett underkänt betyg stadigt ökar från årskurs 6 till årskurs 12. Omkring hälften av eleverna med underkänt betyg i årskurs 6 når godkända betyg i årskurs 9. En större andel elever som i årskurs 6 inte hade några underkända betyg fick underkända betyg i årskurs 9. Ett liknande mönster ses vid övergången från årskurs 9 till årskurs 12. En större andel elever gick från att inte ha underkända betyg i årskurs 9 till minst ett underkänt betyg i årskurs 12, jämfört med andelen elever som kunde förbättra sina underkända betyg från årskurs 9 till att enbart få godkända betyg i årskurs 12.

finns elever med utländsk bakgrund (Källa: GOLD).



En närmare granskning av undergrupperna av elever (figur 5) visar att elever med svensk bakgrund uppvisar samma mönster av en ökande andel elever med minst ett underkänt betyg över årskurserna 6 till 12. Andelen elever med utländsk bakgrund med minst ett underkänt betyg är högre i årskurs 6 och 9 jämfört med elever med svensk bakgrund. Den stora ökningen av andelen elever med utländsk bakgrund med minst ett underkänt betyg från årskurs 6 till årskurs 9, vänds till största delen vid övergången från årskurs 9 till årskurs 12. Till årskurs 12 är det liten skillnad mellan elever med svensk och utländsk bakgrund för andelen elever med minst ett underkänt betyg. Detta kan vara en återspeglning av initiala språksvårigheter i samband med ökande svårighetsgrad av ämnesinnehåll, vilket senare vänds när svenska språkkunskaper upphör att vara ett hinder för elevers lärande. Figur 5 visar också att även om det är en något högre andel pojkar med både svensk och utländsk bakgrund än flickor med minst ett underkänt betyg så är mönstret för pojkar och flickor likartat.

Förskjutna betygsskala – ekvidistant och kontinuerlig betygsskala

Inledningsvis tog vi upp att betygsskalan är konstruerad med en absolut gräns för betyget E samt att betyget F ges en särskilt stor vikt vid översättning till poäng, eftersom avståndet mellan betygen F och E är 10 poäng, jämfört med 2,5 poäng mellan övriga betygssteg. Att få

betyget F ger således en kraftfullt negativ effekt för meritvärdet. Detta kan jämföras med att ge 27,5 poäng för betyget A, vilket ger 10 poängs skillnad mellan betygen B och A, och därmed en kraftfull positiv effekt för elever med betyget A. För att undersöka hur en betygsskala som är ekvidistant (dvs. har lika stora poängavstånd mellan de olika betygsstegen) och kontinuerlig (dvs. saknar en absolut gräns) påverkar meritvärdet och rankningen, har vi kodat om F till 7,5 poäng, så att betygsskalan blir ekvidistant och kontinuerlig från 7,5–20 poäng.

Med hjälp av denna omkodning kan vi undersöka vad konsekvenserna blir för meritvärdet för alla elever, och för olika grupper av elever, hur rankningen påverkas på nationell nivå (dvs. för alla elever i en årskurs), samt jämföra utfallet med den ordinarie betygsskalan, där betyget F ger 0 poäng,

I Tabell 4 presenteras jämförelser av meritvärdet och jämförelsetalet vid omkodningen där F är 0 respektive 7,5. Antalet underkända betyg är av stor betydelse. Skillnaderna vid omkodningen är inte avsevärda för elever med ett till två underkända betyg, medan elever med tre eller fler underkända betyg får en ökning av meritvärdet och jämförelsetalet med cirka 48 och 46 poäng i respektive årskurs 6 och 9, samt 1,9 poängs ökning av jämförelsetalet i årskurs 12.

Tabell 4. Jämförelse av meritvärde (åk 6, 9) och jämförelsetal (åk 12) för F=0 och F=7,5 i årskurs 6, 9 och 12 för 1, 2 eller 3 ämnen.

	Åk 6		Åk 9		Åk 12	
	F = 0	F = 7,5	F = 0	F = 7,5	F = 0	F = 7,5
0 F	225		241		15,2	
1 F	173	180	178	181	12,8	13,1
2 F	149	163	145	157	11,7	12,3
3 F	95	143	87	133	9,0	10,9

Källa: GOLD.

I Tabell 5 presenteras resultat av omkodningen och hur detta påverkar meritvärdet och jämförelsetalet för tre underkända betyg i de behörighetsgivande ämnena svenska/Sva, matematik och engelska.

Tabell 5. Jämförelse av meritvärde (åk 6, 9) och jämförelsetal (åk 12) för F = 0 och F = 7,5 i årskurs 6, 9 och 12 för svenska/Sva, matematik och engelska (behörighetsgivande ämnena).

	Åk 6		Åk 9		Åk 12	
	F = 0	F = 7,5	F = 0	F = 7,5	F = 0	F = 7,5
Meritvärde (åk 6, 9), jämförelsetal åk 12	73	135	50	115	7,81	10,42

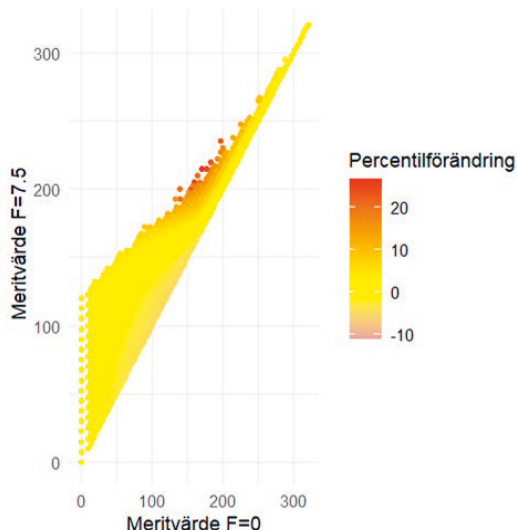
Källa: GOLD.

I resultaten framkommer att effekten av omkodningen från F = 0 till F = 7,5 poäng påverkar rankningen i mycket liten omfattning. I genomsnitt förändras rankningen med 0,008 % i årskurs 6 och 0,05 % i årskurs 9. Förändringen sker främst i nedre delen av distributionen, där elever som fått F=0 poäng och som i omkodningen fått 7,5 poäng. Att ge F = 7,5 poäng innebär små förändringar i rankningen för lågpresterande elever, medan rankningen för övriga elever är densamma och de behåller sin position på rankningen.

Att ge betyget F 7,5 poäng leder således *inte* till några formella konsekvenser eller effekter på betygen som urvalsinstrument. Däremot kan elevernas tilltro till sin kompetens och förmåga och motivation påverkas positivt om de får poäng även för kunskaper under E. Att betyget F renderar viss poäng är dessutom rimligt, då elever som deltagit i undervisning har vissa kunskaper i ämnet, även om de inte når upp till samtliga kriterier för betyget E.

För att tydligare visa hur omkodningen slår avseende rankningen presenteras dessa resultat grafiskt.

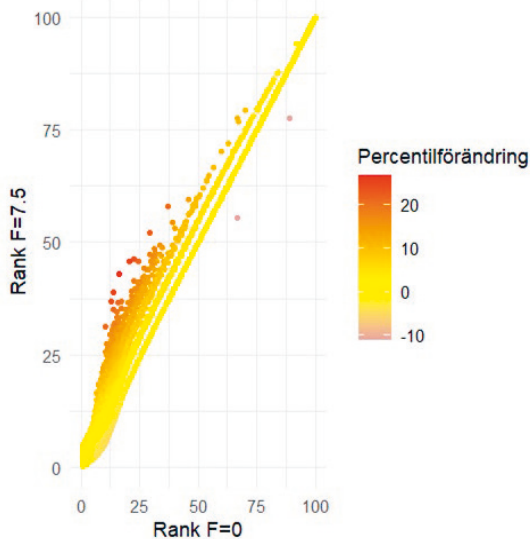
Figur 6. Meritvärde före och efter tilldelning av 7,5 poäng för underkänt betyg, storleken på förändringen i percentilgraden (Källa: GOLD).



Figur 6 visar meritvärdena före och efter tilldelning av 7,5 poäng för underkänt betyg. Färgen representerar hur mycket deras percentilrankning ändrades – röd representerar elever med högre rang efter att ha tilldelats 7,5 poäng till ett F, medan lila representerar elever med lägre rang efter förändringen, och gult representerar ingen förändring i rang. Spridningen av poängen i den nedre änden av meritvärdeskalen indikerar att individer med de lägsta meritvärdena kommer att öka sina poäng mest. Detta mönster minskar med ökande meritvärde, och det sker inte någon förändring alls för eleverna nära den högsta änden av meritvärdeskalen. Med hänvisning till förändringen i percentilrankningen är det ett fåtal individer runt mitten av meritvärdeskalen som får viss förändring i rang, och flera som marginellt ökar sin percentil runt den mellersta tredjedelen av skalan. Inga elever drabbades av kraftiga nedgångar i rang efter att ha tilldelats 7,5 poäng för underkänt betyg. I stället visar vissa elever på den nedre halvan av meritvärdeskalen små nedgångar i sin percentilgrad. Det är elever vars meritvärde inte har ökat

med det ändrade värdet på underkänt betyg, förmodligen för att de inte har många underkända betyg. Ingen elevs meritvärde blev mindre efter förändringen, vilket framgår av den tydliga raka linjen på den nedre sidan av diagrammen.

Figur 7. Percentilrankning före och efter tilldelning av 7,5 poäng till ett underkänt betyg, storleken på förändringen i percentilrankningen. (Källa: GOLD).



Figur 7 visar hur varje elevs percentilrankning förändras om F-betygen tilldelades 7,5 poäng. Individer som uppvisar en stor förbättring i sin percentilrankning är röda, medan de med lägre ställning efter förändringen är färgade lila. De bästa resultaten kommer att synas närmare percentilrankning 100, medan de lägsta resultaten finns närmare percentilrankningen 0. Diagrammen visar att de största ökningarna i percentilrang är mellan ungefär den 10:e och 40:e percentilen, medan mindre ökningarna i percentilen förekommer över nästan hela skalan. Majoriteten av individer som visar en lägre percentilrankning efter att ha tilldelat 7,5 poäng till ett underkänt betyg tenderar att finnas i den nedre delen av skalan. De gula markeringarna i figuren är individer som inte har ändrat sin rang efter tilldelning av poäng för underkänt betyg. Inom ett betyg, t.ex. E, så finns ett stort antal elever som i viss mån byter plats på rankingen inom detta betygssteg. Det är få som byter rang över betygssteg.

Kapitel 4: Slutsatser

I detta kapitel sammanfattar vi rapportens resultat i relation till tidigare forskning och dataanalyserna.

Som beskrivs i kapitel 1 är det inte alltid tydligt i tidigare forskning om det finns en absolut godkänträn i studier om konsekvenser av betygssättning. Få studier undersöker specifikt effekterna av en sådan gräns. Mycket av det vi vet om betygens pedagogiska effekter kommer således från studier om låga betyg, men också från olika betygssystem. Det finns emellertid inget i denna forskning som tyder på att närvaron av en absolut godkänträn skulle påverka dessa konsekvenser i en positiv riktning, t.ex. öka elevernas incitament att anstränga sig, eftersom de negativa konsekvenserna av ett underkänt betyg är mer kännbara. Som diskuteras i kapitel 2 finns det snarare en risk att flera av de negativa effekterna förstärks, eftersom det är mer stigmatiserande och mer känslomässigt nedbrytande att få ett underkänt betyg jämfört med ett lågt.

Som framgått ovan, finns det inte något tydligt stöd i tidigare forskning för att en absolut godkänträn skulle ge en positiv effekt på elevers motivation eller prestationer. Det är också tveksamt i vilken utsträckning det faktiskt går att fastställa, med någon form av rimlig precision, vilka elever som faktiskt når samtliga kriterier för betyget E och vilka som inte gör det, vilket ytterligare försvagar de argument som framförts för nyttan av en sådan gräns; som att det skulle leda till ökad motivation hos elever och vara ansvarsutkrävande för föräldrar, lärare och skolor. I graferna som presenterats i kapitel 3 framkommer dessutom att elever med underkända betyg i hög grad fortsätter att få underkända betyg senare i skolgången, vilket motsäger hypotesen och förhoppningen att F ska leda till en långsiktig minskning av F.

- Att ge 7,5 poäng för F innebär minskade skillnader i meritvärdet mellan elever som fått F och de som fått E men innebär inte en betydelsebärande skillnad för rankningen och betygen som urvalsinstrument. F skulle fortsatt vara det lägsta betyget och ge minst antal poäng. Att förändra den nuvarande betygsskalan till en *ekvidistant skala* vore önskvärt då en sådan skala mättekniskt är mer korrekt än den nuvarande betygsskalan och de negativa konsekvenserna är obefintliga för utbildningssystemet medan de positiva konsekvenserna är stora för den enskilde eleven.
- Betydelse för likvärdigheten är att skalan bör vara kontinuerlig, vilket innebär att den absoluta gränsen för E tas bort. I flera utvärderingar framkommer att lärare tolkar betygsriterier för den absoluta gränsen (E) olika. Att elevers kunskaper inte bedöms likvärdigt av alla lärare är ett hot mot principen om likvärdiga betyg och ska därför inte förekomma. Om den absoluta gränsen tas bort innebär det att lärare betygssätter elevers kunskaper på en glidande skala från F–A från 7,5–20 poäng, där F är det lägsta betyget och A det högsta.
- Betyget F i nuvarande betygssystem innebär att elever inte nått upp till betygsriterierna för betyget E. Betyget F innebär *inte* att eleven inte har några kunskaper i ämnet utan det är troligt att elever som har deltagit i undervisningen har vissa kunskaper. Därför är det rimligt att vissa kunskaper under E bör rendera en viss poäng. Att låta F motsvara t.ex. 7,5 poäng skulle innebära ett något högre meritvärde för de allra mest lågpresterande men har inga konsekvenser för den totala rankningen. Däremot kan det innebära att elever med flera underkända betyg blir motiverade av att

få något högre poäng, vilket kan påverka deras motivation positivt och ge dem möjlighet att utveckla positiva lärandestrategier.

- Om kunskaper under E innebär viss poäng skulle differentieringen mot streck bli tydligare.

För att tydliggöra så innebär omkodning av F till 7,5 poäng att det finns två alternativ: 1) F görs om till det lägsta betygssteget och ger 7,5 poäng eller 2) F tas bort och i stället utgörs betygsskalan av E – A. I båda dessa scenarier tas det underkända betygssteget bort och ersätts av ett lägsta betygssteg som ger den lägsta poängen. Betygskriterier blir i viss mån utan betydelse för det lägsta betygssteget men kan skrivas fram som en grundläggande nivå som alla elever klarar som går i svensk skola.

Oavsiktligt utfall av underkända betyg

Den stora andelen elever med underkända betyg kan beskrivas som ett överraskande och oavsiktligt utfall av införandet av ett icke godkänt betygssteg. Överlag har diskussioner uteblivit som problematiserat de kriterierelaterade betygen och införandet av en godkäntgräns (Peters, 1996). En specificerad lägsta godkänd nivå har över tid i stället setts som en nödvändighet för att få elever att lära sig, vilket bygger på en logik att det är incitament som belöningar och bestraffningar som driver elever att lära. Betyg är en valuta som elever ska drivas av oavsett individuella förutsättningar och ambitioner. Skolmisslyckanden i svensk grundskola institutionaliserades i form av ett nytt uttryckligt underkänt betyg (Arensmeier, 2022a).

Förväntningen på att en godkäntgräns skulle göra skolan till en bättre plats för lågpresterande elever har inte infriats. Tvärtom, mönstret för en betydande andel av elever som misslyckas kvarstår och är slående stabilt (Arensmeier, 2022a; 2022b). Både skolor och elever har, på grund av förändringarna i betygslogik och språk, blivit misslyckanden (jfr Ball, 2015; Tiana, 2008; Wikström, 2006).

Vad i stället?

Mot bakgrund av forskning inom bedömningsområdet och de empiriska analyser som denna rapport presenterat är det rimligt att:

1. förändra betygsskalan så att den blir ekvidistant och kontinuerlig genom att ge F 7,5 poäng och låta betygskriteriet för E omfatta F, alternativt
2. ta bort F ur betygsskalan och ha en ekvidistant och kontinuerlig betygsskala med enbart godkända betyg som går från t.ex. E till A där E = 10, D = 12,5, C = 15, B = 17,5 och A = 20.

Riskerna med en sådan förändring är små när det gäller hur ett sådant betygssystem skulle fungera som urvalsinstrument. Det kan däremot få följder med avseende på hur betygskriterier kan användas och hur de skrivs fram. Det kan vara önskvärt att ha betygskriterier eller någon typ av beskrivning av vilka kunskaper som är betydelsefulla och som omfattar den lägsta betygsnivån.

De positiva effekterna av en förändring av betygsskalan skulle vara avsevärda eftersom elever med låga och underkända betyg får poäng, vilket är en betydelsefull skillnad mot att få ingenting (0 poäng). Detta bör kunna påverka elevernas känslomässiga reaktioner som t.ex. ökad tilltro till sin förmåga och ökad motivation att lära.

Slutord

Oavsett ideologisk och politisk tillhörighet bör det vara av intresse för Sverige att konstruera ett betygssystem med så få negativa konsekvenser som möjligt för elevernas lärande och utveckling och samtidigt säkerställa ett så rättvist urvalssystem som möjligt. Att förändra betygsskalan genom att kunskaper under E antingen renderar poäng eller F tas bort ur betygsskalan kommer att innebära att betygsskalan blir ekvidistant och kontinuerlig, vilket innebär att det finns låga till höga betyg som ger lärare möjlighet och förutsättningar att göra mer likvärdiga bedömningar.

Referenser

- Arensmeier, C. (2022a). Institutionalizing school failure: From abandoning to reintroducing a failing grade — the rationales behind Swedish grading reforms. *Journal of educational change*, 23 (2), 221–252.
- Arensmeier, C. (2022b). Three Decades of School Failure in Swedish Compulsory School. *Scandinavian Journal of Educational Research*, 66 (1), 14–27.
- Ball, S. J. (2015). What is policy? 21 years later: reflections on the possibilities of policy research. *Discourse: Studies in the Cultural Politics of Education*, 36(3), 306–313.
- Black, P., Harrison, C., Lee, C., Marshall, B. & Wiliam, D. (2004). Working inside the black box: Assessment for learning in the classroom. *Phi Delta Kappan*, 86, 8–21.
- Butler, R. (1988). Enhancing and undermining intrinsic motivation: The effects of task-involving and ego-involving evaluation on interest and performance. *British Journal of Educational Psychology*, 58, 1–14.
- Collins, M. & Lundstedt, J. (2021). The long-term consequences of more informative grading. *Lund Papers in Economic Demography, 2021:1*. Ekonomihögskolan, Lunds universitet.
- Harlen, W. & Deakin-Crick, R. (2002). *A systematic review of the impact of summative assessment and tests on students' motivation for learning*. London: EPPI-Centre, Social Science Research Unit, University of London Institute of Education.
- Hirsh, Å. (2020). When assessment is a constant companion: students' experiences of instruction in an era of intensified assessment focus. *Nordic Journal of Studies in Educational Policy*, 6, 89–102.
- Högberg, B., & Strandh, M. (2024): Temporal trends and inequalities in school-related stress in three cohorts in compulsory school in Sweden. *Scandinavian Journal of Educational Research*.

- Högberg, B., Lindgren, J., Johansson, K., Strandh, M. & Petersen, S. (2021) Consequences of school grading systems on adolescent health: evidence from a Swedish school reform, *Journal of Education Policy*, 36(1), 84-106.
- Jönsson, A. (2022). Perceptions of assessment: An interview study of participants' perceptions of being assessed in Swedish adult education colleges. *Frontiers in Education: Assessment, Testing and Applied Measurement*, 7(836334).
- Jönsson, A. & Balan, A. (2018). Analytic or holistic: A study of agreement between different grading models. *Practical Assessment, Research & Evaluation*, 23(12).
- Jönsson, A., Balan, A. & Hartell, E. (2021). Analytic or holistic grading? *Assessment in Education: Principles, Policy & Practice*, 28, 212–227.
- Kefer, E. (1975). Relationships between academic achievement and personality characteristics: a quasi-longitudinal design. *American Educational Research Journal*, 12, 191–210
- Klapp, A. (2015). Does grading affect educational attainment? A longitudinal study. *Assessment in Education: Principles, Policy and Practice*, 22(3), 302-323.
- Koenka, A. C., Linnenbrink Garcia, L., Moshontz, H., Atkinson, K. M., Sanchez, C. E. & Cooper, H. (2019). A metaanalysis on the impact of grades and comments on academic motivation and achievement: a case for written feedback. *Educational Psychology*, 41, 922–947.
- Koretz, D. (2008). *Measuring up: What educational testing really tells us*. Harvard, MA: Harvard University Press.
- Linder, A., Gerdtham, U-G. & Heckley, G. (2023). Child mental health: Impact of introducing earlier compulsory school grades. *Working Paper 2023:2*. Ekonomihögskolan, Lunds universitet.
- Löfgren, H., Alm, F., Jönsson, A., Hultén, M., Markström A-M., & Lundahl, C. (2021). *Betyg i årskurs 4: En studie om bedömningspraktikerna på skolor som deltagit i försöksverksamhet med tidigare betyg*. Stockholm: Skolverket.
- Löfgren, R. & Löfgren, H. (2016). *Att få sina första betyg: En rapport om elevers berättelser om sina erfarenheter av att få betyg i årskurs 6*. Stockholm: Skolverket.
- Pekrun, R., Frenzel, A. C., Goetz, T. & Perry, R. P. (2007). The control-value theory of achievement emotions: An integrative approach to emotions in education. In P. A. Schutz & R. Pekrun (Eds.), *Emotion in education* (pp. 13–36). Elsevier Academic Press.
- Pérez Prieto, H. & Löfgren, H. (2017). *Att ständigt bli bedömd: Elevers berättelser om betyg och nationella prov*. Lund: Studentlitteratur.

- Peters, B. G. (1996). Political institutions, old and new. In R. E. Goodin & H.-D. Klingemann (red.), *A new handbook of political science* (pp. 205–220). Oxford: Oxford University Press.
- Rönn, C. & Petterson, D. (2023). Swedish students' everyday school life and teachers' assessment dilemmas: Peer strategies for ameliorating schoolwork for assessment. *Educational Assessment, Evaluation and Accountability*, 35, 37–66.
- Sadler, R. (2005) Interpretations of criteria-based assessment and grading in higher education. *Assessment and Evaluation in Higher Education*, 30, 175–194.
- Sadler, D. R. (2009). Indeterminacy in the use of preset criteria for assessment and grading. *Assessment & Evaluation in Higher Education*, 34(2), 159–179.
- Skolverket (2015). *Skolreformer i praktiken: Hur reformerna landade i grundskolans vardag 2011-2014*. Stockholm: Skolverket.
- Skolverket (2024). *Bedömning för hållbart lärande*. Stockholm: Skolverket.
- SOU 2016:25 *Likvärdigt, rättssäkert och effektivt – Ett nytt nationellt system för kunskapsbedömning*. Stockholm: Utbildningsdepartementet.
- Tiana, A. (2008). Should 'failing' students repeat a grade/the year? *Journal of Educational Change*, 9(4), 391–394.
- Vogt, B. (2017). *Just assessment in school: A context-sensitive comparative study of pupils' conceptions in Sweden and Germany*. Doktorsavhandling, Linnéuniversitetet.
- Wahlström, N. & Sundberg, D. (2015). En teoribaserad utvärdering av läroplanen Lgr 11. *Rapport 2015:7*. Uppsala: IFAU.
- Wikström, C. (2006). Education and assessment in Sweden. *Assessment in Education: Principles, Policy & Practice*, 13(1), 113–128.

Betydelsen av en skarp godkäntgräns i betygssystemet

Anna Sjögren, Institutet för arbetsmarknads- och utbildningspolitisk utvärdering (IFAU), Uppsala Center for Labor Studies (UCLS)

Jonas Vlachos, Stockholms universitet

Björn Öckert Institutet för arbetsmarknads- och utbildningspolitisk utvärdering (IFAU), Uppsala universitet, Uppsala Center for Labor Studies (UCLS)

2024-10-02

Sammanfattning

Rapporten analyserar konsekvenserna av en skarp godkäntgräns i betygssystemet, med fokus på den svenska skolan. Enligt internationella jämförelser har Sverige inte en särskilt hög andel lågpresterande elever, men andelen underkända elever är hög. Utrymmet för att minska andelen underkända elever genom generella förbättringar av utbildningens kvalitet torde därför vara begränsad. En mindre godtycklig bedömning av elevers kunskaper – exempelvis genom extern rättning av nationella prov och en koppling av betygen till dessa prov – skulle troligen öka andelen underkända elever markant, givet att inga andra åtgärder vidtas. Internationell forskning tyder på att skarpa godkäntgränser under vissa omständigheter kan ge starkare incitament för förbättrade resultat, främst då för elever nära gränsen. Forskningen tyder emellertid även på att dylika gränser kan ge upphov till kontraproduktivt strategiskt beteende bland skolor och elever. Gränser kan också verka demotiverande, främst bland elever i socialt utsatta miljöer. Forskningen finner också tydliga negativa effekter av att underkännas på framtida utbildning, arbetsmarknadsutfall och kriminalitet. Däremot saknas tydliga belägg för att generösare antagningsregler skulle försämra utbildningens kvalitet.

Innehållsförteckning

1	Introduktion	3
2	Vad är grundproblemet?	4
3	Förväntade effekter av en skarp godkänthgräns?	12
4	Beskrivande empirisk evidens för betydelsen av en godkänthgräns för skolresultat och progression i Sverige	16
5	Kausala studier och evidens från andra länder	32
5.1	Studier som mäter incitamenteffekter av ökade krav	33
5.2	Konsekvenser av att missa kunskapsmålen	36
5.3	Samlade effekter av att införa kunskapskrav	39
6	Slutsatser.....	42
	Referenser	44
	Bilaga	49

1 Introduktion

Målet för skolväsendet är att ge alla elever så goda kunskaper och färdigheter som möjligt för att de ska kunna växa upp till välfungerande vuxna.¹ Ett sätt att uppnå detta är att sträva efter att alla elever uppnår en viss miniminivå av kunskaper och färdigheter. Men frågan uppstår; var ska nivån ligga, hur ska den mätas, vad händer om nivån inte uppnås och hur påverkar den undervisningen och lärandet för elever med olika förutsättningar?

I grundskolan och gymnasieskolan bedöms elevers kunskaper och färdigheter genom betyg som sätts av lärare, med stöd av nationella prov (tidigare standard- och centralprov). Betygen fungerar både som motivation och information till eleverna, men utgör också urvalsinstrument till nästa nivå i utbildningssystemet. I den funktionen är det centralt att betygen är rättvisande för att fördelningen av utbildningsplatser ska vara både rättvis och effektiv.

Den svenska skolan har haft en skarp gräns för godkända betyg sedan införandet av det målrelaterade betygssystemet (Lpo 94 och Lpf 94). I en målstyrd skola kan denna gräns fungera som ett verktyg för ansvarsutkrävande, där skolor och lärare ska säkerställa att elever lämnar skolan med en acceptabel nivå av kunskaper (Figlio och Loeb 2011). I grundskolan är godkänthöjden avgörande för behörighet till gymnasieskolans nationella program och i gymnasieskolan påverkar den möjligheterna att uppnå grundläggande behörighet till högskolan.²

Sedan godkänthöjden infördes har dess betydelse och konsekvenser varit föremål för diskussion och debatt. Å ena sidan har det påpekats att andelen elever som inte blir behöriga till gymnasieskolans nationella program – och därmed utestängs från fortsatt yrkesutbildning – har legat stadigt mellan 10 och 15 procent, vilket grundskolan inte har lyckats åtgärda under de senaste 30 åren. Å andra sidan har det framhållits att elever med svaga förkunskaper från grundskolan har svårt att klara kraven i gymnasieskolan, vilket gör att sänkta krav vore olyckligt. Dessutom finns en oro för att ett starkt fokus på godkänthöjden påverkar hela undervisningen negativt. Det saknas dock en

¹ I skollagens Kap1:4 uttrycks det så här ” Barn och elever ska ges stöd och stimulans så att de utvecklas så långt som möjligt. En strävan ska vara att uppväga skillnader i barnens och elevernas förutsättningar att tillgodogöra sig utbildningen. Utbildningen syftar också till att i samarbete med hemmen främja barns och elevers allsidiga personliga utveckling till aktiva, kreativa, kompetenta och ansvarskännande individer och medborgare.”

² Idag gäller att eleven måste ha godkända betyg i kärnämnen svenska, matematik, och engelska samt ytterligare fem ämnen för att komma in på ett yrkesprogram och kärnämnen plus ytterligare nio ämnen för att komma in på ett studieförberedande program. Dessa ämnen ska dessutom inkludera programmets karaktärsämnen.

systematisk genomgång av den forskning som berör konsekvenserna av en godkänträn för elever och skolsystemet som helhet.

Huvudsyftet med denna rapport att ge en översikt av det aktuella kunskapsläget om hur elevers skolgång påverkas av en skarp grän för godkända betyg, med fokus på den svenska skolan. Vi kommer att belysa effekter på skolresultat samt arbetsmarknadsutvecklingen för de elever som påverkas. Särskild uppmärksamhet kommer att riktas mot följderna av att inte uppnå behörighet till gymnasieskolans nationella program. En annan målsättning är också att diskutera möjliga förklaringar till de observerade effekterna. Sammanställningen kommer i första hand avse forskningsresultat inom utbildningsekonomi, och viss uppdatering av resultat ifrån flera tidigare IFAU-rapporter (Holmlund m.fl. 2019; Holmlund m.fl. 2014; Björklund m.fl. 2010). Vi kommer även i viss mån att beröra relevant forskning inom andra discipliner.

Vi inleder med en problembeskrivning som undersöker förekomsten av elever med låga kunskapsresultat i Sverige i jämförelse med andra länder. Därefter går vi utifrån ekonomisk teori igenom tidigare studier av hur elever, lärare, skolor och skolsystem kan förväntas påverkas av en skarp godkänträn av det slag som finns i Sverige. Vi redogör också för hur utvecklingen av deltagande och genomströmning i gymnasieskolan sett ut sedan det tidiga 1990-talet, då den skarpa godkäntränen infördes. Vi sammanfattar även resultatet av svensk och internationell forskning som på olika sätt berör de kausala effekterna av godkäntränser eller behörighetsregler innan vi sammanfattar våra slutsatser.

2 Vad är grundproblemet?

Att elever har olika förutsättningar att uppnå samma studieresultat borde vara okontroversiellt, liksom att det är svårt att fastställa en absolut lägsta nivå för de kunskaper och färdigheter som unga människor behöver ha med sig ut i livet. Exakt hur elevers förutsättningar att lära är fördelade i befolkningen är inte helt klarlagt, men de är troligen ungefär normalfördelade.³ Detta har implikationer för elevernas studieresultat och för vilka följder en skarp godkänträn får: Oavsett samhällets ambitioner för skolan, behöver man beakta konsekvenserna av en lägstanivå i betygssystemet, och de därtill kopplade behörighetsreglerna, för såväl de svagaste elevernas förutsättningar för egen försörjning och ett gott liv, som för övriga elever. I detta avsnitt diskuteras fördelningen av elevers resultat och vilka elever som har svårt att nå upp till OECD:s eller den svenska

³ Detta var t.ex. utgångspunkten i utformningen av de standardproven. Se Husén 1956

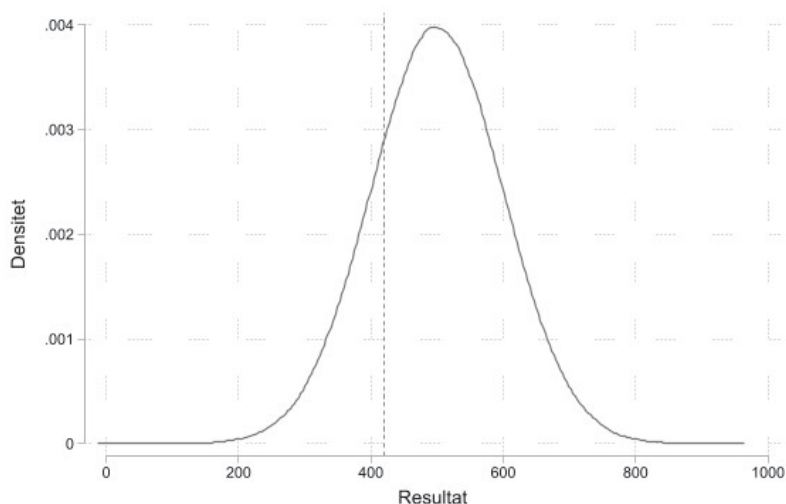
skolans mål. Vi analyserar resultat från PISA-undersökningen, de nationella proven samt betygen.

Med en normalfördelning avses en fördelning som är symmetrisk kring ett medelvärde, där cirka två tredjedelar av observationerna befinner sig inom en standardavvikelse över respektive under detta medelvärde. Vidare befinner sig 95 procent av observationerna inom två standardavvikelser från medelvärdet, och i princip alla – 99,7 procent av observationerna – hamnar inom tre standardavvikelser från medelvärdet. Det finns en stor forskningslitteratur om fördelningen av kognitiva förmågor. Litteraturen är alltför omfattande för att redovisas i detalj här, men det finns en bred enighet om att dessa förmågor är ungefärligen normalfördelade.⁴ Den vetenskapliga diskussionen gäller huvudsakligen huruvida extrema avvikelser – både uppåt och nedåt – är vanligare än normalfördelningen låter påskina. För syftet med denna rapport är denna aspekt av marginell betydelse.

Studieresultat påverkas inte enbart av kognitiv förmåga utan även av faktorer som motivation, ansträngning, stöd och intresse. Enligt den centrala gränsvärdessatsen är dock summan av dessa faktorer normalfördelad, förutsatt att de är oberoende. Även om antagandet om oberoende troligen inte är uppfyllt, är studieresultat i praktiken ungefärligen normalfördelade. Figur 1 illustrerar betydelsen av en skarp godkäntgräns när studieresultat är normalfördelade. I figuren antas ett genomsnittresultat på 500 och en standardavvikelse på 100, vilket motsvarar fördelningen av PISA-resultaten. OECD, som ansvarar för PISA-undersökningen, har även fastställt en miniminivå av kunskaper (nivå 2) som alla elever bör uppnå. För matematik motsvarar denna nivå 420 poäng. Elever som inte uppnår minimikraven befinner sig till vänster om den vertikala linjen i figuren.

⁴ Det finns en omfattande forskningslitteratur kring kognitiva förmågor och deras fördelning. Se t.ex. följande översikter: Jensen (1998); Flynn (2007); Hunt (2010); Nisbett m.fl. (2012); Deary (2020).

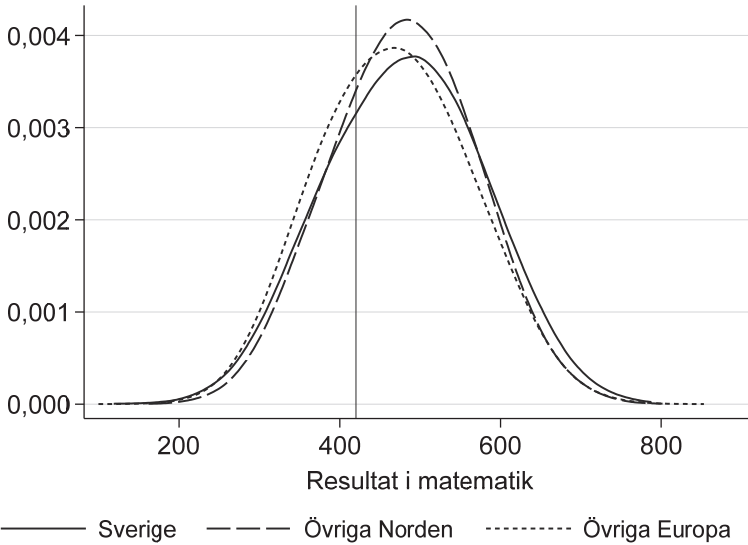
Figur 1 Illustration av fördelning med gräns vid 420 (matematik nivå 2 enligt PISA)



Not: Simulerade resultat utifrån ett medelvärde på 500 och en standardavvikelse på 100. Den vertikala linjen anger 420, motsvarande OECD:s nivå 2.

Utifrån de antaganden som gjorts om fördelningen innebär detta att vi teoretiskt kan förvänta oss att 21 procent av eleverna inte uppnår de ämneskunskaper som OECD anser vara tillräckliga. Denna förutsägelse ligger nära vad som faktiskt observeras. Figur 2 visar resultatfördelningen i matematik enligt PISA-undersökningen 2022 för Sverige, övriga Norden samt övriga Europa. Även om fördelningarna inte är identiska ligger de nära varandra, och visar att en betydande andel av eleverna inte når upp till OECD:s miniminivå. Likartade mönster står att finna även i övriga PISA-ämnen (Figur A1, se bilaga). Det bör dock noteras att OECD:s definition av tillräckliga ämneskunskaper framstår som relativt godtycklig och vad som anses vara "tillräckliga" kunskaper rimligen beror på det sammanhang som avses.

Figur 2 Resultatfördelning i matematik för olika länder (PISA 2022)



Not: Egna analyser av data från PISA 2022. Den vertikala linjen indikerar gränsen för nivå 2.

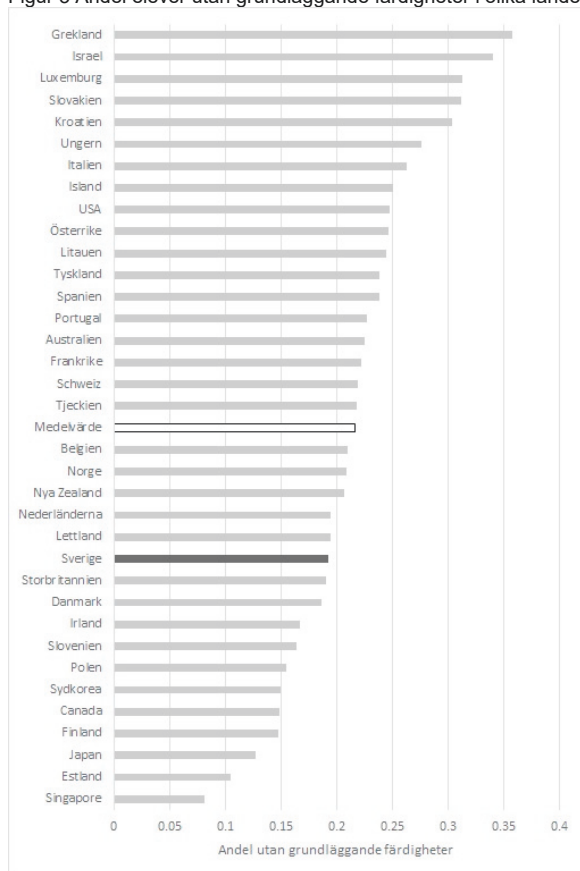
Att en betydande andel elever inte når upp till miniminivån är tydligt också i en studie av Gust m.fl. (2024). De analyserar elevprestationerna i matematik och naturorienterade ämnen från de internationella PISA- och TIMSS-undersökningarna och översätter resultaten till en gemensam skala. Figur 3 visar andel elever som inte uppnår grundläggande färdigheter, vilket uppgår till i genomsnitt 22 procent i det urval av länder som studeras. I Sverige är andelen 19 procent, medan Singapore är landet med lägst andel: åtta procent. Det bör noteras att Singapore uppnår exceptionellt goda resultat i internationella undersökningar, och inget annat land har en andel under 10 procent. I andra högprestera länder som Estland, Japan och Sydkorea är andelen elever utan grundläggande färdigheter 10, 13 respektive 15 procent.⁵

Detta är av betydelse då det ger en indikation om det realistiska utrymmet för att via förbättrade ämneskunskaper få en större andel att klara den uppställda miniminivån. Det finns alltid utrymme för förbättrade resultat men det framstår

⁵ Det bör noteras att det i många länder uppstår tidvis infekterade diskussioner om huruvida det egna landets resultat under- eller överskattas. Så även i Sverige (Riksrevisionen 2021). Poängen här är att inget land har en försurbar andel elever som inte lyckas.

som osannolikt att en väsentligt lägre andel än 15 procent av svenska elever skulle uppnå detta. Om vi utgår från antagandet om normalfördelning skulle detta kräva att den genomsnittliga prestationsnivån höjs till nivån i några av världens mest högpresterande länder. Även vid en sådan förbättring kommer således en betydande andel av eleverna fortfarande inte ha uppnått den nivå som OECD anger som godtagbar.

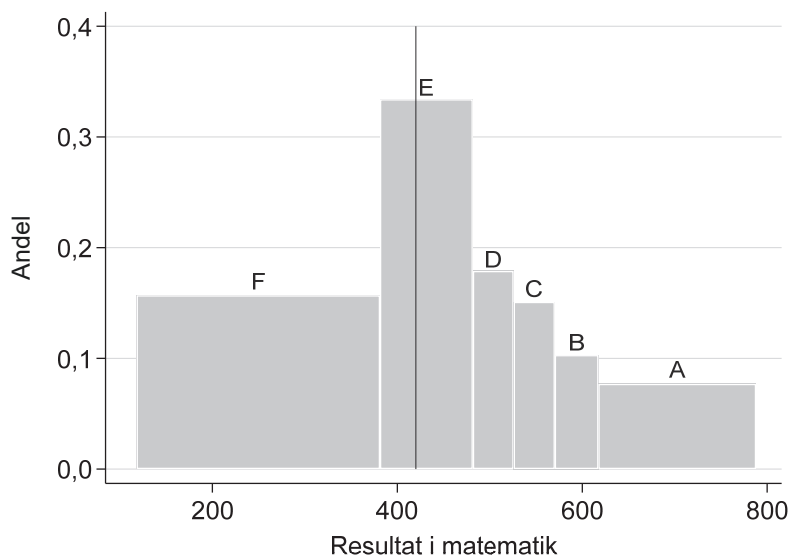
Figur 3 Andel elever utan grundläggande färdigheter i olika länder



Källa: Gust m.fl. (2024).

En relevant fråga är hur OECD:s kravnivå förhåller sig till kravnivån i den svenska skolan. Figur 4 visar fördelningen av resultaten på de nationella proven i matematik i årskurs 9 för läsåret 2021/22, där resultaten översatts till den skala som används i PISA-undersökningen. I figuren anges andel elever som uppnår olika betyg på de nationella proven och vilket resultat motsvarande andel elever skulle ha haft i PISA-undersökningen 2022. Den vertikala linjen markerar gränsen för att uppnå OECD:s prestationsnivå 2, vilket 27 procent av de svenska PISA-deltagarna inte gjorde. Samtidigt var det omkring 15,7 procent av de elever som skrev det nationella provet som inte uppnådde godkänt resultat, dvs. lägst betyget E. Eftersom svårighetsgraden på de nationella proven i matematik varierar mellan åren finns en betydande variation i andelen som inte når godkänd nivå. Läsåret 2018/19 uppnådde 14,5 procent av de skrivande eleverna inte godkänt resultat, medan andelen var 10,9 procent läsåret 2022/23 (läsåren 2019/20 och 2020/21 ställdes de nationella proven in på grund av pandemin).⁶

Figur 4 Resultat på de nationella proven i matematik i årskurs 9 för läsåret 2021/22

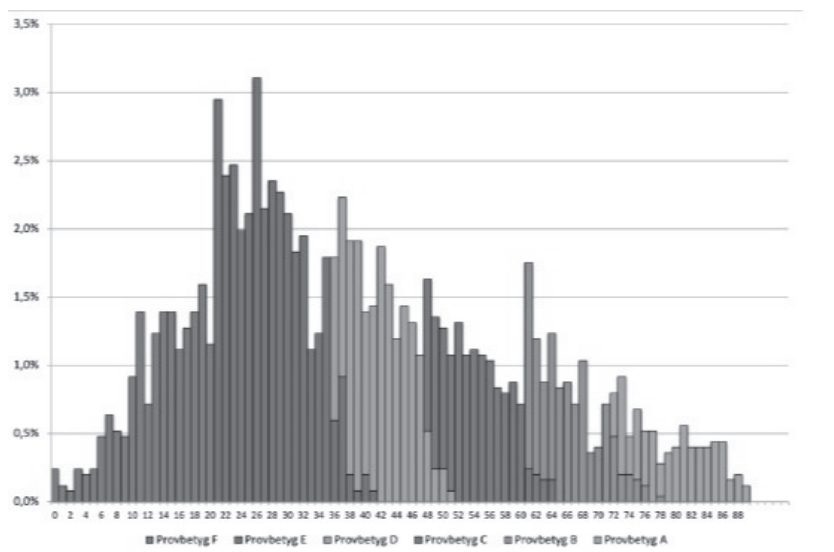


Källa: Skolverket, egen bearbetning.

⁶ Elever kan ha starka incitament att prestera väl på de nationella proven, vilket kan bidra till att en större andel klarar de nationella proven jämfört med andelen som uppnår nivå 2 i PISA-undersökningen.

Förutom problemet med bristande stabilitet i de nationella provens svårighetsgrad finns det flera felkällor som gör att andelen elever som uppnår godkänt betyg på de nationella proven överskattas. En sådan felkälla är att proven rättas lokalt på skolorna och inte av en oberoende instans. Som illustreras i Figur 5 finns en tendens bland rättande lärare att justera upp resultaten för elever som befinner sig nära en betygsgräns; det finns färre elever än förväntat precis under betygsgränsen och motsvarande en högre andel precis över denna gräns. Denna tendens är särskilt tydlig för elever nära E-nivån, vilket medför att andelen som klarar godkänd nivå är högre än den hade varit vid extern bedömning. Detta är inget unikt svenskt fenomen utan något som observerats på många håll (t.ex. Dee m.fl. (2019)).

Figur 5 Poängfördelning för de nationella proven i matematik 2021/22



Källa: Ek (2022).

Ett annat problem är att inte alla elever skriver de nationella proven. Läsåret 2021/22 avslutade 119 570 elever matematikkursen i årskurs 9, men endast 104 247 (87,8 procent) deltog i de nationella proven. Analyser av data från TIMSS 2015 visar att lågpresterande elever är överrepresenterade i gruppen som saknar provresultat. Av de 17 procent av eleverna som inte skrev det nationella provet i matematik i årkurs 9 läsåret 2015/16 hade drygt 35 procent ett resultat

på TIMSS-provet i matematik som motsvarar betyget F.⁷ Om man antar att provbetygen skulle ha fördelat sig på samma sätt även för de elever som inte deltog i det nationella provet i matematik läsåret 2021/22, skulle andelen elever som inte klarar godkänthöjden vara ungefär 18 procent. Motsvarande justering skulle medföra att andelen elever som inte når godkända resultat i matematik skulle vara 19 procent för läsåret 2018/19 och 15 procent för läsåret 2022/23.

Slutligen finns en tydlig tendens att elever får högre slutbetyg än provbetyg. Denna tendens är särskilt markant i matematik, speciellt för elever som inte uppnått godkänd nivå på de nationella proven. Läsåret 2021/22 fick, som tidigare nämnts, 15,7 procent av eleverna betyget F på det nationella provet i matematik. Andelen elever som samma läsår fick slutbetyget F i detta ämne var dock bara 10,3 procent. Sammantaget samverkar dessa faktorer till att överdriva andelen elever som uppfyller kraven för godkänd nivå i matematik i den nationella betygsstatistiken.

Detta avsnitt visar att kunskapsresultaten i Sverige, liksom i andra länder, följer en ungefärlig normalfördelning. Jämfört med likartade länder utmärker sig inte Sverige vad gäller andelen lågpresterande elever. Detta innebär att andelen elever som riskerar att hamna under en given kunskapsnivå beror på var man lägger ribban. Om det anses önskvärt med en stipulerad lägstanivå för elevernas kunskaper finns det i princip bara två sätt att åtgärda problemet med elever som inte klarar kraven. Det ena är en generell förbättring av alla elevers kunskaper – att flytta upp hela resultatfördelningen – och det andra är att sänka kraven för vad som anses vara en acceptabel lägstanivå.

Analyserna av de internationella kunskapsjämförelserna visar att det framstår som realistiskt att via generella kunskapsförbättringar påverka andelen lågpresterande elever så att problemet blir försumbart. Även världens mest högpresterande skolsystem har en betydande andel elever som presterar dåligt. Det alternativ som kvarstår är således att sänka nivån för vad som anses vara en godkänd nivå, inte minst för att uppnå behörighet till gymnasieskolans yrkesprogram.

Som framgår av detta avsnitt sker i praktiken redan en sänkning av kunskapskraven i Sverige, då flera faktorer samverkar för en mer generös bedömning av elevernas kunskaper. Det sker dock på ett godtyckligt sätt där provens svårighetsgrad, skolors och lärares bedömningar av provresultat, samt provens inverkan på betygssättningen spelar stor roll. En viktig slutsats är därför att en opartisk bedömning av de nationella proven och en tydligare förankring

⁷ TIMSS-resultaten i matematik 2015 för de 17 procent av eleverna som inte deltog i det nationella provet i matematik 2015/16 fördelar sig enligt följande provbetyg: A: 4,27 procent, B: 6,72 procent, C: 9,37 procent, D: 11,22 procent, E: 32,55 procent samt F: 35,87 procent.

av betygen mot dessa provresultat, med bibehållna krav för godkänt, skulle resultera i en markant ökning av andelen elever som underkänns i grundskolan. Det skulle medföra att fler elever nekas behörighet till gymnasieskolans nationella program, och därmed till en utbildning som ger förutsättningar för att de ska kunna försörja sig själva som vuxna.

3 Förväntade effekter av en skarp godkäntgräns?

Vad säger ekonomisk teori och tidigare empirisk forskning om hur en skarp gräns för godkända betyg, behörighet till gymnasieskolan eller examination, påverkar kunskapsutvecklingen i skolan? För att besvara denna fråga behöver man fundera över hur godkäntgränsen påverkar såväl elever som lärare och skolor. Vilka konsekvenser följer av att elever når eller inte når ett godkänt betyg och vilka möjligheter har elever, lärare och skolor att agera för att påverka måluppfyllelsen? Eftersom många aspekter av betygssystem och antagnings-system samverkar, är det svårt att renodla diskussionen till att enbart gälla förekomsten av en skarp godkäntgräns. Därför för vi diskussionen utifrån de förutsättningar som gäller i en svensk kontext. Diskussionen utgår bl.a. från en omfattande litteratur som undersökt betydelsen av examinationskrav och betygsnivåer, samt politik som syftar till ansvarsutkrävande i skolsystemet (*accountability*) baserat på standardiserade tester och lägstanivåer för kunskaper, se t.ex. (Becker och Rosen 1992; Jacob 2005; Bowden, m.fl. 2023; Neal 2010; Neal och Schanzenbach 2010; Deming och Figlio 2016; Figlio och Loeb 2011; Mansfield och Slichter 2024).

I Sverige skiljer sig konsekvenserna av godkända betyg mellan olika utbildningsnivåer. I grundskolan innebär godkända slutbetyg i kärnämnen (svenska eller svenska som andraspråk, matematik och engelska), samt i minst fem ytterligare ämnen, att en elev blir behörig till gymnasiets nationella yrkesprogram.⁸ För gymnasieelever krävs godkända betyg i kärnämnen, samt i minst 90 procent av kurserna, för att erhålla en gymnasieexamen⁹ och att uppnå grundläggande behörighet till högre studier.

Underkända betyg begränsar alltså möjligheterna att studera vidare och att konkurrera om jobb. I grundskolan räcker det med underkänt i svenska, matematik eller engelska för att en elev inte ska få tillträde till ett nationellt program. Underkända betyg drar också ner meritvärdet kraftigt, eftersom

⁸ För behörighet till studieförberedande program krävs godkänt i kärnämnen samt i minst ytterligare nio ämnen.

⁹ Därtill krävs godkänt gymnasiearbete samt, för yrkesprogram, minst 400 godkända poäng i de programgemensamma ämnena.

underkända betyg ges värdet noll. Skillnaden mellan F och E är 10 poäng, medan det skiljer 2,5 poäng mellan övriga betygssteg. Det innebär att underkända betyg inte bara påverkar behörigheten, utan också försämrar konkurrenskraften för studieplatser, även för dem som är behöriga till ett nationellt program.

Konsekvenserna av att få underkänt i ett ämne är kännbara för elever. Elever som ligger på gränsen för att uppnå behörighet till gymnasiet, få sin gymnasieexamen eller nå behörighet till högre studier har starka incitament att anstränga sig i skolan. Enligt ekonomisk teori kommer kraven för att nå ett visst betyg, eller för att uppnå godkänt betyg i ett ämne, att påverka elevernas ansträngningar att lära sig. Om ansträngningarna är ändamålsenliga, påverkar de även deras lärande (Becker och Rosen 1992). Förutsatt att elever strävar efter godkända betyg, kommer deras ansträngningar att öka när kraven höjs, så länge de vet hur de ska förbättra sina resultat och bedömer att ansträngningen ökar deras chanser att nå målet.¹⁰

För elever som befinner sig långt över gränsen för godkänt, kommer en skärpning av kraven inte att göra någon skillnad. Elever som redan tidigare kämpade för att klara sig kan däremot ge upp om uppgiften upplevs för svår, dvs. om sannolikheten att lyckas är för liten i förhållande till hur mycket ansträngningen kostar. Teoretiskt kan skarpa godkäntgränser alltså förväntas stimulera till ökade ansträngningar bland de elevgrupper som befinner sig på gränsen till godkänt, där en rimlig ansträngning ökar chansen att bli behörig. Högpresterande elevers ansträngningar påverkas inte av godkäntgränsen utan snarare av gränser för, och konkurrens om, högre betyg, och att upprätthålla sin position i resultatfördelningen.

För elever som upplever att det är svårt eller omöjligt att nå upp till godkäntgränsen kan den istället vara demotiverande, inte bara i det ämne där eleven fallerar utan också i andra ämnen, eftersom det räcker med underkänt betyg i ett kärnämne för att förlora möjligheten att studera vidare. Utifrån ett psykologiskt perspektiv kan denna process beskrivas som att underkända betyg påverkar elevers motivation negativt genom att skada deras tilltro till sin egen förmåga att lära sig, och genom att få dem att känna att de inte hör hemma i skolan (Schwartz m.fl. 2016). Ett starkt betygsfokus i skolan kan också riskera att tränga undan elevers inneboende motivation och lust att lära (Ryan och Deci 2000).

Även skolor och lärare påverkas av förekomsten av godkäntgränser och kan på olika sätt anpassa betygssättning, undervisning och organisation. Lärare vill oftast göra det bästa för sina elever och står dessutom under tryck från

¹⁰ Experimentella studier visa att elever svarar på incitament, men att de inte nödvändigtvis vet hur de ska nå bättre studieresultat (Fryer 2011; Levitt m.fl. 2016).

huvudmän, samhället och föräldrar att stötta eleverna så att de uppnår godkända betyg. Skolor blir jämförda och utvärderade av huvudmän och politiker samt av föräldrar som väljer skola. Huvudmän och samhället tenderar att använda enkla mått på kvalitet, såsom andelen elever som når målen och det genomsnittliga meritvärdet, vilka publiceras på Skolverkets hemsida. Som beskrivits tidigare har underkända elever en betydande negativ effekt på skolans snittvärden. En mer sofistikerad metod är att också ta hänsyn till skillnader i skolornas elevsammansättning. Skolverkets Salsa-mått gör detta. Föräldrar bryr sig också om dessa kvalitetsmått, men är särskilt intresserade av hur skolan kan främja deras eget barns utveckling samt om barnet skulle trivas och passa in.

Lärare och skolor påverkas alltså av skarpa godkäntgränser, men konsekvenserna för betygssättning, undervisning och organisation beror på vilka medel som står till buds och hur enkelt eller kostsamt det är att påverka elevernas betyg. En övergripande slutsats från litteraturen om *accountability*-system är att skolväsendet, på gott och ont, svarar mot de incitament som finns i systemet. Jacob (2005) visar t.ex. att lärare i Chicago fokuserade undervisningen på ämnesområden och moment som testades, och att fler elever undantogs från proven när skolor började utvärderas på basis av testresultat. Det skapas alltså drivkrafter för att lägga tid och undervisningsresurser på att lyfta resultaten i de ämnen och områden som utvärderas eller får störst genomslag i utvärderingen. Beroende på hur lärare och skolor anpassar sig kommer också konsekvenserna för elevers lärande att variera.

Utän normerande externa bedömningar av elevernas kunskaper blir det för lärare och skolor enklast att undvika underkända elever genom att tolka betygskriterier flexibelt och sätta generösa betyg, vilket i sin tur leder till betygsinflation. Denna drivkraft begränsas dock sannolikt av lärares och rektorers professionalism, eftersom det strider mot läroplaner och regelverk. Samtidigt visar Figur 5 tydlig evidens på att elever precis under en skarp poänggräns tenderar att bedömas mer generöst eller rättas upp. Det finns också tydliga tecken på att betygssättningen är mer generös i ämnen utan nationella prov än i ämnen med prov. De nationella proven har alltså, trots den generösa rättningen, en normerande effekt. I linje med detta har betygssättningen också blivit mer generös under pandemin, när nationella prov inte genomfördes, se t.ex. Svaleryd och Vlachos (2021).

Beroende på om sänkta krav i rättning och betygssättning begränsas till endast godkäntgränsen eller även påverkar de högre betygen, kan vissa eller alla elevers lärande påverkas, vilket diskuterats i detta avsnitt. Det innebär att vissa elevers lärande kan gynnas medan andras missgynnas. En följd av generös betygssättning är att elever vars utbildningsval annars hade begränsats av

godkänthöjningen gynnas eftersom de får möjlighet att studera vidare, även om deras faktiska kunskapsnivå kanske är lägre. Lärare och skolor kan också anpassa undervisningen på olika sätt. Med en begränsad budget kan förekomsten av godkänthöjningar leda till att större vikt läggs vid de ämnen och moment som utgör de största hindren för att elever ska nå godkända betyg, medan mindre vikt läggs vid ämnen och moment som de flesta elever klarar av. Undervisningen kan även anpassas till de elever som riskerar att få underkänt, vilket kan leda till en nivåsänkning och minskad stimulans för övriga elever. Resurser för särskilt stöd kan också riktas till elever som har möjlighet att lyftas över godkänthöjningen i årskurs 9, eller i slutet av gymnasiekursen, snarare än till elever i lägre årskurser eller till elever som har svårt att nå godkänt trots stöd (Neal och Schanzenbach 2010). Trots att behovet av tidigt stöd har diskuterats länge, är det fortfarande vanligare att elever får stöd i årskurs 9 än i lågstadiet (Holmlund m.fl. 2019).

Ett annat sätt för skolor att undvika att elever får underkända betyg är att påverka skolans elevunderlag. En skarp godkänthöjning kan driva på skolsegregation om skolor har möjlighet att påverka sin elevsammansättning. För fristående skolor är detta möjligt genom att anpassa inriktning, marknadsföring och undervisning så att skolan blir ett attraktivt alternativ för elever som förväntas klara målen. Kommunala huvudmän, som behöver erbjuda alla elever skolplats, kan välja att sprida ut svaga elever i olika skolor eller samla dem på vissa skolor, beroende på hur de arbetar med dessa elever. I grundskolan, där fördelningen av elever styrs av skolval och närhetsprincipen, är utrymmet begränsat, även om inriktningar som musik och idrott kan påverka elevsammansättningen. Skolval kan också bidra till segregering om familjer väljer skolor som förväntas rikta undervisningen mot deras barns specifika behov. I gymnasieskolan, med betygsintag och olika program, påverkas andelen elever som riskerar att inte klara godkänthöjningen av konkurrensen om studieplatser och programutbudet.

På systemnivå kan en skarp godkänthöjning skapa ett tryck på skolmyndigheter och politiker att tillhandahålla verktyg för att minska andelen underkända elever, eller att mildra konsekvenserna för en skola eller huvudman att ha elever som riskerar att bli underkända. I avsaknad av resurser till särskilt stöd eller undervisning, som kan lyfta dessa elever, kan efterfrågan i stället styras mot ökade möjligheter att få undantag från vissa krav eller rätt till olika hjälpmedel för vissa elever.

4 Beskrivande empirisk evidens för betydelsen av en godkänthgräns för skolresultat och progression i Sverige

Det är svårt att empiriskt belägga kausala effekter av skarpa godkänthgränser eftersom de vanligen införs för alla elever på en gång, vilket innebär att en lämplig jämförelsegrupp saknas. Eftersom dessa betygsgränser främst berör elever som riskerar att befinna sig nära den nya gränsen, är det möjligt att undersöka grupper av elever som kan förväntas påverkas olika mycket av införandet av en ny gräns. Denna ansats används bl.a av Holmlund m.fl. (2019), som beskriver hur andelen elever som påbörjar och avslutar gymnasiet har utvecklats över tid. Studien visar hur reformer i gymnasieskolan, men också förändringar i betygssystemet i grund- och gymnasieskolan, har påverkat elever i olika delar av resultatfördelningen. I detta avsnitt följer vi denna metod samt uppdaterar och vidareutvecklar analysen. Vi begränsar dock beskrivningen till elever som är födda i Sverige eller som invandrat före skolstart.

Flera förändringar har skett i gymnasieskolan sedan tidigt 1990-tal. De tvååriga gymnasieprogrammen förlängdes med ett år och utökades med teoretiskt innehåll för att ge högskolebehörighet. Från läsåret 1992/93 infördes också ett individuellt program för elever som inte sökt eller inte kommit in på det gymnasieprogram de sökt. Avsikten var att fånga upp alla elever i gymnasieåldern under en tid när arbetsmarknadsutsikterna för dem försämrades, delvis till följd av 1990-talskrisen. Därtill förändrades betygssystemet med införandet av Lpo94 i grundskolan och Lpf 94 i gymnasiet, vilket innebar att eleverna fick målrelaterade bokstavs betyg (G, VG, MVG) istället för de tidigare relativa betygen (1–5). I gymnasiet övergick man från att sätta ett slutbetyg i varje ämne till att sätta betyg på olika kurser inom samma ämne, vissa redan under det första året.

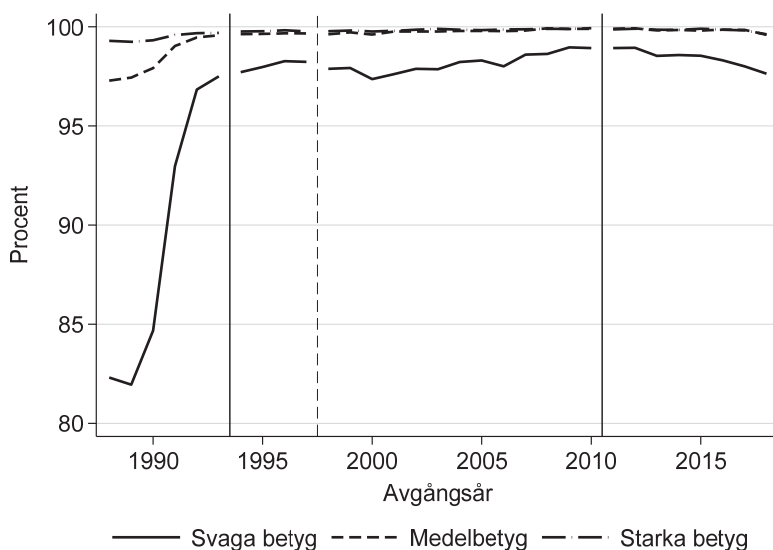
Införandet av betyget godkänt i grundskolan innebar att godkända betyg i kärnämnen svenska, matematik och engelska blev behörighetskrav för att komma in på gymnasiets nationella program. Detta innebar en betydande skärpning av behörighetskraven, då det tidigare endast krävdes avgångsbetyg från grundskolan för att få tillträde till yrkes- och högskoleförberedande linjer på gymnasiet.¹¹ I och med införandet av Gy11 skärptes behörighetskraven ytterligare. För att bli behörig till yrkesförberedande program krävdes godkänt betyg i kärnämnen och ytterligare fem ämnen, medan det för studieförberedande program krävdes godkänt i kärnämnen och ytterligare nio

¹¹ För behörighet till en del teoretiska linjer krävdes dock att eleven hade betyg i vissa ämnen samt att eleven hade läst särskild kurs i engelska eller matematik (Skolöverstyrelsen 1984).

ämnen. Dessutom förändrades betygsskalan från G, VG, MVG till fem olika godkända betygsnivåer (A–E).

Figur 6 Andel som börjar på gymnasiet, uppdelat på olika grundskolebetyg visar andelen av avgångselever från grundskolan som påbörjar gymnasiet före 19 års ålder, uppdelat på deras betyg från grundskolan. Eleverna har delats in i tre lika stora grupper: de med högst, medel respektive lägst medelbetyg (meritpoäng). Figurens vertikala linjer visar, i tur och ordning, (1) införandet av målrelaterade betyg på gymnasiet, (2) den första kohort som berörs av målrelaterade betyg i grundskolan och därmed skärpta behörighetsregler för gymnasiet (streckad linje), och (3) införandet av Gy11 med ytterligare skärpning av behörighetskraven. Figur 6 Andel som börjar på gymnasiet, uppdelat på olika grundskolebetyg visar att i stort sett alla grundskoleelever idag går vidare till gymnasiet, att den kraftiga expansionen av gymnasiet i början av 1990-talet främst berörde svagpresterande elever och att det sedan dess finns en liten grupp av elever med svaga betyg som inte går vidare. Men vad gör dessa elever på gymnasiet?

Figur 6 Andel som börjar på gymnasiet, uppdelat på olika grundskolebetyg

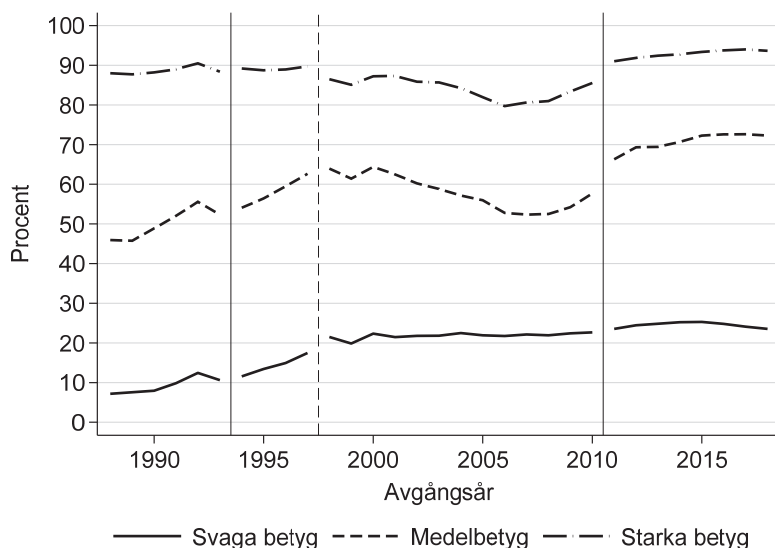


Not: Figuren visar andel avgångselever som registrerats på gymnasieskolans första år före 19 års ålder, för den tredjedel av eleverna med de sämsta grundskolebetygen, tredjedelen med medelbetyg, samt tredjedelen med de bästa betygen. Den första

heldragna lodräta linjen anger införandet av det mål- och kunskapsrelaterade betygssystemet på gymnasiet, den streckade linjen visar införandet av behörighetskrav till nationella program medan den andra heldragna linjen anger införandet av Gyll.

Figur 7 och Figur 8 visar andelen elever som påbörjar högskoleförberedande respektive yrkesförberedande linje/program. Även om i stort sett alla elever börjar på gymnasiet sedan gymnasieexpansionen i början av 1990-talet, så har det skett förändringar i vad eleverna studerar. De flesta elever med höga betyg, omkring 80–90 procent, studerar på högskoleförberedande program. Andelen minskade dock marginellt när behörighetskrav till gymnasieskolan infördes 1998. I samband med Gyll – då yrkesprogrammen inte längre gav grundläggande behörighet till högskolan – valde elever med höga betyg i större utsträckning högskoleförberedande program.

Figur 7 Andel som börjar på högskoleförberedande linje/program på gymnasiet, uppdelat på olika grundskolebetyg

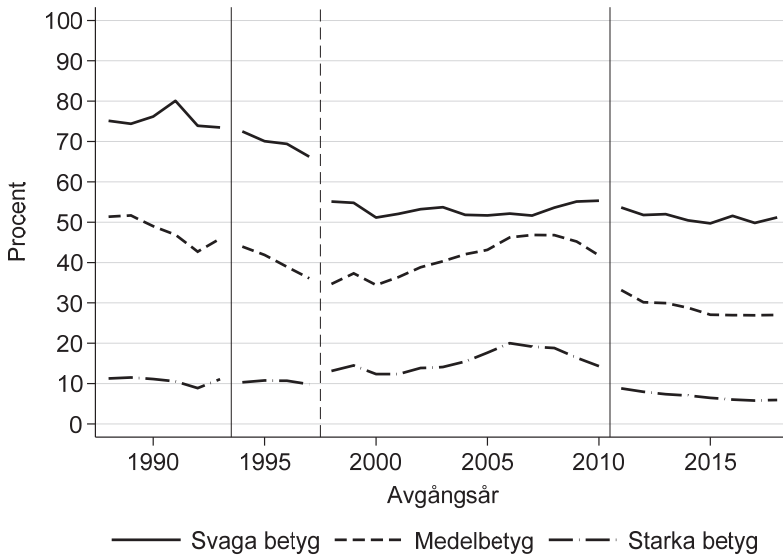


Not: Figuren visar andel avgångselever som börjar på högskoleförberedande linjer/program före 19 års ålder, för den tredjedelen av eleverna med de sämsta grundskolebetygen, tredjedelen med medelbetyg, samt tredjedelen med de bästa betygen. Den första heldragna lodräta linjen anger införandet av det mål- och kunskapsrelaterade betygssystemet på gymnasiet, den streckade linjen visar införandet av behörighetskrav till nationella program medan den andra heldragna linjen anger införandet av Gyll.

Andel elever med genomsnittliga betyg som går på högskoleförberedande program har ökat till över 70 procent i slutet av perioden, jämfört med knappt 50 procent i början av 1990-talet. Val av yrkesprogram har minskat i motsvarande grad. En markant ökning av elever som väljer studieförberedande program – över 10 procentenheter – skedde i samband med Gyll-reformen, vilket motsvaras av en minskning av elever på yrkesprogrammen.

Bland elever med svaga betyg har andelen på högskoleförberedande program också ökat över tid, från under 10 procent av kohorten till omkring 25 procent. Samtidigt har andelen på yrkesprogram minskat från mellan 70 och 80 procent till omkring 50 procent. När behörighetskraven till gymnasiet införs, väljer något fler lågpresterande elever högskoleförberedande program, medan färre väljer yrkesprogram. Vid införandet av Gyll sker också i denna grupp en omfördelning av elever från yrkesprogram till högskoleförberedande program.

Figur 8 Andel som börjar på yrkeslinje/program på gymnasiet, uppdelat på olika grundskolebetyg

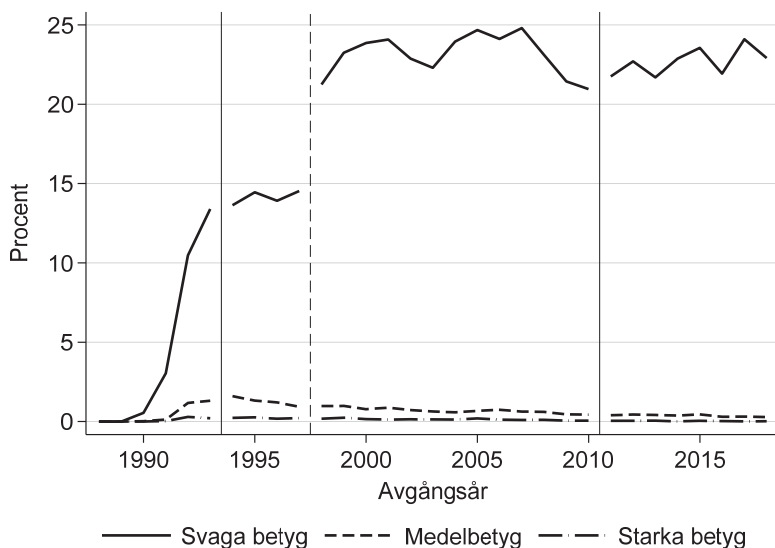


Not: Figuren visar andel avgångselever som börjar på yrkeslinjer/program före 19 års ålder, för den tredjedelen av eleverna med de sämsta grundskolebetygen, tredjedelen med medelbetyg, samt tredjedelen med de bästa betygen. Den första heldragna lodräta linjen anger införandet av det mål- och kunskapsrelaterade betygssystemet på gymnasiet, den

streckade linjen visar införandet av behörighetskrav till nationella program medan den andra heldragna linjen anger införandet av Gyll.

Den kanske mest dramatiska förändringen under perioden är ökningen av andelen elever som går på individuella/introduktionsprogram, vilket framgår av Figur 9. Utvecklingen – som inleddes med introduktionen av det individuella programmet 1992/93 – gäller dock bara elever med låga grundskolebetyg. I samband med skärpningen av behörighetskraven till gymnasiet 1998/99 ökade andelen studiesvaga elever på det individuella programmet med drygt sex procentenheter. Sedan dess går nästan var fjärde elev med låga grundskolebetyg på individuella/introduktionsprogram. Vid gymnasiereformen 2011 skärptes behörighetskraven till de nationella programmen – vilket framförallt påverkade studiesvaga elever – och andelen på introduktionsprogrammen ökade med omkring två procentenheter i denna grupp.

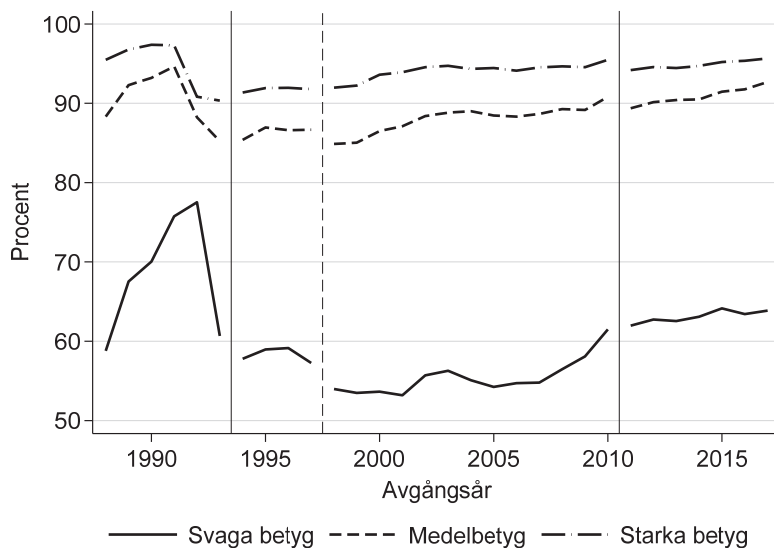
Figur 9 Andel som börjar på individuella/introduktionsprogram på gymnasiet, uppdelat på olika grundskolebetyg



Not: Figuren visar andel avgångselever som börjar på individuella/introduktionsprogram före 19 års ålder, för den tredjedelen av eleverna med de sämsta grundskolebetygen, tredjedelen med medelbetyg, samt tredjedelen med de bästa betygen. Den första heldragna lodräta linjen anger införandet av det mål- och kunskapsrelaterade

betygssystemet på gymnasiet, den streckade linjen visar införandet av behörighetskrav till nationella program medan den andra heldragna linjen anger införandet av Gy11.

Figur 10 Andel av en årskull som avslutat gymnasiet vid 19 års ålder, uppdelat på grundskolebetyg



Not: Figuren visar andel avgångselever som avslutat gymnasiet vid 19 års ålder, för den tredjedelen av eleverna med de sämsta grundskolebetygen, tredjedelen med medelbetyg, samt tredjedelen med de bästa betygen. Elever antas ha fullföljt sina gymnasiestudier om de antingen erhållit slutbetyg eller varit registrerade på det sista året av ett nationellt program. Den första heldragna lodräta linjen anger införandet av det mål- och kunskapsrelaterade betygssystemet på gymnasiet, den streckade linjen visar införandet av behörighetskrav till nationella program medan den andra heldragna linjen anger införandet av Gy11.

Ett motiv bakom de skärpta behörighetsreglerna är att elever behöver vissa förkunskaper för att klara gymnasiet, och att införandet av det individuella programmet skulle hjälpa studiesvaga elever att fullfölja sina studier. Utvecklingen tyder dock inte på att så blev fallet. Figur 10 visar andel av en årskull som avslutat gymnasiet vid 19 års ålder, alltså på normaltid, uppdelat i

de tre grupperna med höga, medel och svaga grundskolebetyg.¹² Den tidiga gymnasieexpansionen – under slutet av 80-talet och början av 90-talet – innebar också att fler elever gick ut. Det är allra tydligast för elever med låga grundskolebetyg, där som mest nästan 80 procent gick ut gymnasiet. Vid införandet av det individuella programmet minskade andelen elever som klarade gymnasiet på normaltids kraftigt. Sannolikt spelar också förlängningen av yrkesprogrammen, med ökat teoretiskt innehåll, en roll för utvecklingen.

Införandet av de målrelaterade kursbetygen på gymnasiet (den första vertikala linjen) innebar att något färre elever klarade gymnasiet, framför allt bland de med svaga grundskolebetyg. När behörighetsreglerna till gymnasieskolans nationella program började gälla (den streckade vertikala linjen) minskade andelen elever som blev klara på normaltids ytterligare. I den svagaste gruppen sjönk andelen som avslutar gymnasiet ned till knappt 55 procent.

Under perioden 1993–98 minskar således andelen av de svagaste eleverna som fullföljer gymnasiet på normaltids med omkring 20 procentenheter. Även om det är svårt att avgöra hur utvecklingen skulle ha sett ut annars är det troligt att minskningen delvis beror på att allt färre elever fick möjlighet att påbörja ett nationellt program. Även införandet av målrelaterade kursbetyg på gymnasiet tycks ha bidragit till försämrade genomströmning, där fler elever hoppar av eller går om en årskurs och färre avslutar gymnasiet. Det är ändå viktigt att påpeka att de flesta, 90 procent av en årskull, går på gymnasiet i tre år, men att andelen som går på introduktionsprogram, snarare än på yrkesprogram, har ökat. Detta innebär att en stor grupp gymnasieungdomar undervisas på grundskolenivå i kärnämnen och därmed inte får tillgång till yrkesutbildning.

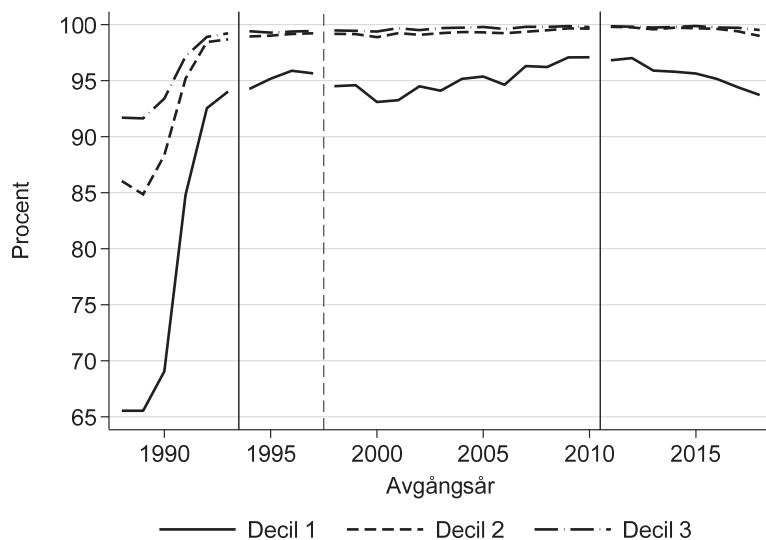
De skärpta behörighetsreglerna påverkade främst elever med svaga grundskolebetyg. För att undersöka om utvecklingen har varit likartad för hela gruppen lågpresterande elever, förfinar vi analysen ytterligare. Vi delar in eleverna i tio lika stora grupper (deciler) utifrån deras grundskolebetyg och studerar de tre lägsta tiondelarna; (1) tiondelen av eleverna med lägsta betyg (decil 1), (2) tiondelen med näst lägsta betyg (decil 2) samt (3) tiondelen med de tredje lägsta betygen (decil 3).

Figur 11 visar andelen elever i de tre lägsta tiondelarna av betygsfördelningen som började läsa på gymnasiet under perioden. Expansionen av gymnasieutbildning i slutet på 1980-talet och i början av 1990-talet tycks främst

¹² Föreskrifterna för betygssättning av elever med hög frånvaro förändrades under mitten av 1990-talet, vilket gör att uppgifter om elever som avslutat gymnasiet är svårtolkade. Vi antar därför att elever har fullföljt sina gymnasiestudier om de antingen erhållit slutbetyg eller varit registrerade på det sista året av ett nationellt program.

ha berott på att elever med de allra svagaste betygen började gå på gymnasiet. Antagligen hade de svårt att hitta jobb under 90-talskrisen. Bland tiondelen elever med de lägsta grundskolebetygen ökade andelen som gick på gymnasiet med omkring 30 procentenheter; från ungefär 65 procent till 95 procent.

Figur 11 Andel svagpresterande elever som påbörjar gymnasiet, uppdelat på betygsdecil



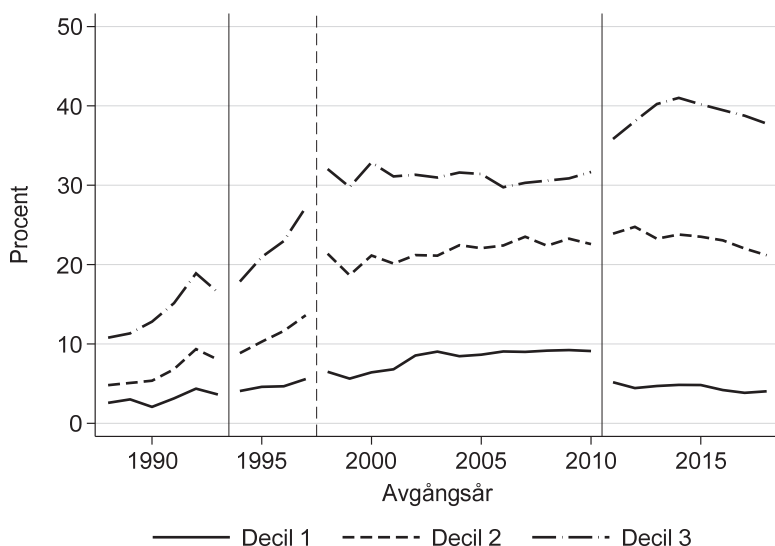
Not: Figuren visar andel avgångselever som registrerats på gymnasieskolans första år före 19 års ålder, för tiondelen av eleverna med de lägsta grundskolebetygen (decil 1), tiondelen med de näst lägsta betygen (decil 2), samt tiondelen med de tredje lägsta betygen (decil 3). Den första heldragna lodräta linjen anger införandet av det mål- och kunskapsrelaterade betygssystemet på gymnasiet, den streckade linjen visar införandet av behörighetskrav till nationella program medan den andra heldragna linjen anger införandet av Gy11.

Figur 12 visar andel lågpresterande elever som går på högskoleförberedande linjer/program, vilket relativt få av dem gör. Det finns dock stora skillnader mellan de olika betygsgrupperna. Bland tiondedelen av elever med de lägsta betygen är det mindre än 10 procent som går på högskoleförberedande linjer/program under hela perioden. För elever med något högre betyg (decil 2

och 3) ökar andelen, framförallt i samband med införandet av de målrelaterade grundskolebetygen 1998 och vid gymnasireformen 2011.

Det är inte helt uppenbart varför införandet av behörighetsreglerna till gymnasiet 1998 påverkade elevernas val av högskoleförberedande program. Möjligen uppfattade elever med svaga, men ändå godkända, betyg från grundskolan att de skulle klara av mer avancerade gymnasiestudier.¹³ Det är också bland eleverna i decil 2, och framför allt i decil 3, som andelen på högskoleförberedande program ökade efter införandet av Gy11. Möjligen berodde det på att dessa elever ville ha möjligheten att kunna läsa vidare i framtiden, eftersom yrkesprogrammen inte längre gav grundläggande behörighet till högskolan.

Figur 12 Andel svagpresterande elever på högskoleförberedande linjer/program, uppdelat på betygsdecil



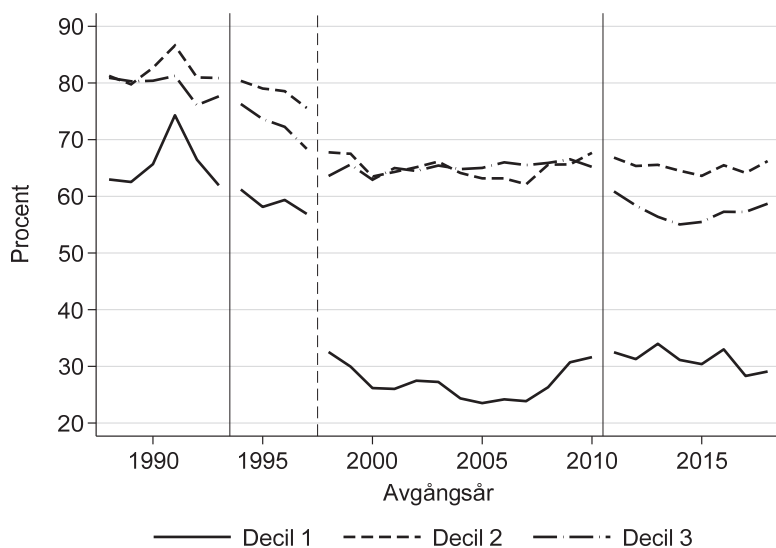
Not: Figuren visar andel avgångselever som börjar på högskoleförberedande linjer/program före 19 års ålder, för tiondelen av eleverna med de lägsta grundskolebetygen (decil 1), tiondelen med de näst lägsta betygen (decil 2), samt tiondelen med de tredje lägsta betygen (decil 3). Den första heldragna lodräta linjen anger

¹³ Man bör emellertid vara försiktig i tolkningarna eftersom grundskolebetygen, som används i rangordningen av eleverna inom varje kohort, inte är helt jämförbara vid byte av betygssystemet.

införandet av det mål- och kunskapsrelaterade betygssystemet på gymnasiet, den streckade linjen visar införandet av behörighetskrav till nationella program medan den andra heldragna linjen anger införandet av Gy11.

Figur 13 visar andelen lågpresterande elever som går på yrkeslinjer/program. Det framgår tydligt att det främst är bland elever med de svagaste betygen (decil 1) som kravet på godkända betyg från grundskolan slår igenom, och andelen som går ett yrkesprogram minskar med 25 procentenheter. Figur 14 visar att behörighetskraven leder till en motsvarande ökning av andelen på individuella program. I betygsdecil 2 och 3 sker ingen ökning av andelen elever på individuella program, vilket tyder på att de är behöriga till nationella program. Minskningen av andelen elever på yrkesprogram bland dessa elever beror istället på att de oftare väljer högskoleförberedande program (Figur 12).

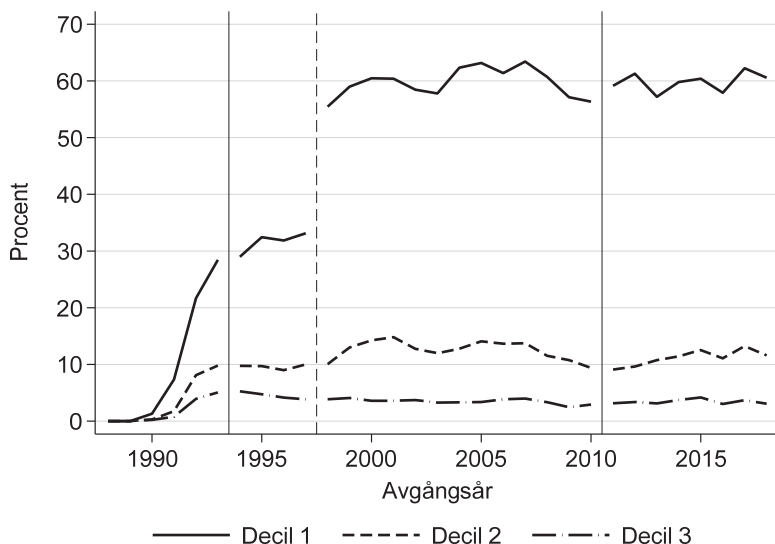
Figur 13 Andel svagpresterande elever på yrkeslinjer/program, uppdelat på betygsdecil



Not: Figuren visar andel avgångselever som börjar på yrkeslinjer/program före 19 års ålder, för tiondelen av eleverna med de lägsta grundskolebetygen (decil 1), tiondelen med de näst lägsta betygen (decil 2), samt tiondelen med de tredje lägsta betygen (decil 3). Den första heldragna lodräta linjen anger införandet av det mål- och kunskapsrelaterade betygssystemet på gymnasiet, den streckade linjen visar införandet

av behörighetskrav till nationella program medan den andra heldragna linjen anger införandet av Gyll.

Figur 14 Andel svagpresterande elever på individuella/introduktionsprogram, uppdelat på betygsdecil



Not: Figuren visar andel avgångselever som börjar på individuella/introduktionsprogram före 19 års ålder, för tiondelen av eleverna med de lägsta grundskolebetygen (decil 1), tiondelen med de näst lägsta betygen (decil 2), samt tiondelen med de tredje lägsta betygen (decil 3). Den första heldragna lodräta linjen anger införandet av det mål- och kunskapsrelaterade betygssystemet på gymnasiet, den streckade linjen visar införandet av behörighetskrav till nationella program medan den andra heldragna linjen anger införandet av Gyll.

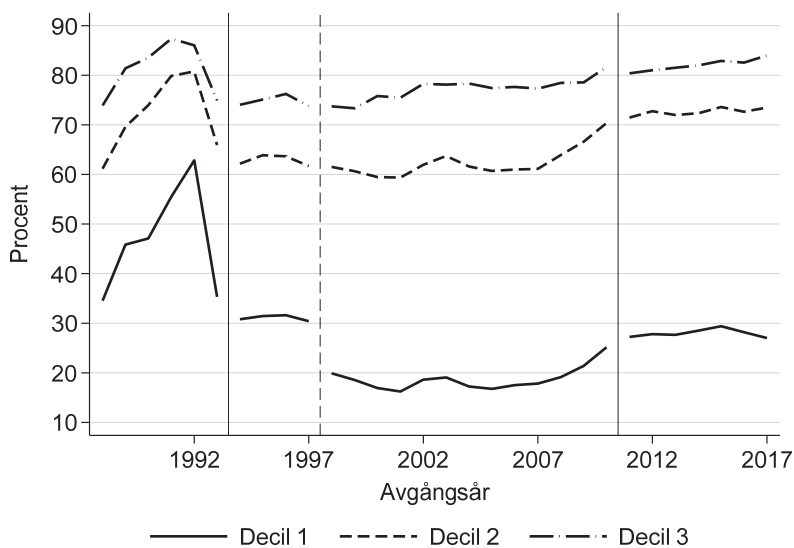
Vi undersöker nu vad som händer med genomströmningen i gymnasieskolan när fler relativt svagpresterande elever går på studieförberedande program och fler riktigt svaga elever går på individuella/introduktionsprogram. Samtidigt förlängdes yrkesprogrammen och fick mer teoretiskt innehåll.

Figur 15 visar att andelen som slutfört gymnasiet vid 19 års ålder ökar i och med gymnasieexpansionen, men minskar efter införandet av det individuella programmet och förlängningen av yrkesprogrammen i början av 1990-talet. Detta syns i alla de tre svagaste betygsdecilerna. Införandet av målrelaterade betyg på gymnasiet ser ut att ha försämrat genomströmningen på gymnasiet något i de två svagaste betygsdecilerna (decil 1 och 2). I samband med införandet av behörighetskrav till gymnasieskolan – och den därtill hörande ökningen av elever på individuella program – minskade andelen som slutförde gymnasieskolan på normalt tid bland elever med de svagaste grundskolebetygen (decil 1). Det ser inte ut som att de skärpta behörighetskraven och förändrade sökbeteendet i samband med Gy2011 har påverkat genomströmningen gymnasieskolan.¹⁴

Figur 16 följer upp eleverna när de är 21 år gamla, då fler har hunnit bli klara med sina gymnasiestudier. Den nedgång i andel elever med de lägsta betygen (decil 1) som fullföljer gymnasiet efter skärpningen av behörighetskraven, som kunde observeras vid 19 års ålder, tycks dock kvarstå. Införandet av en skarp godkänthetsgräns i grundskolan – med därtill hörande behörighetsregler – verkar således ha försämrat genomströmningen på gymnasiet för de svagaste eleverna även på lite längre sikt.

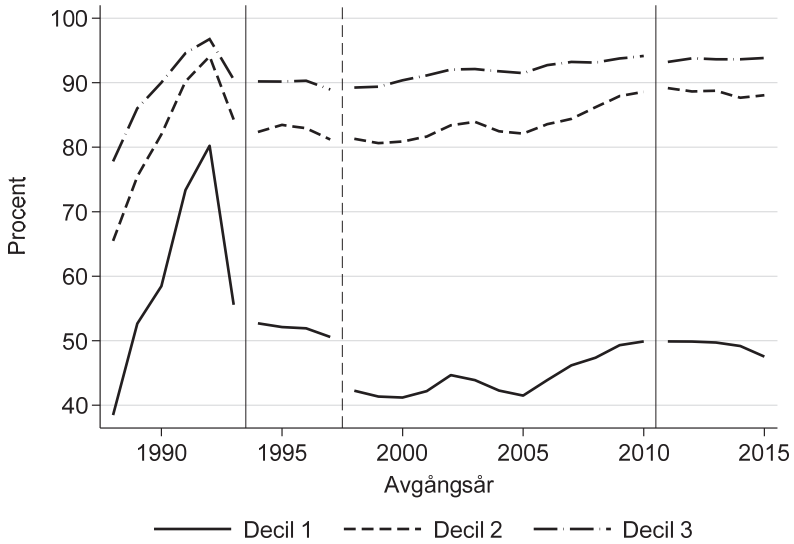
¹⁴ Det var mycket få elever som påverkades av de skärpta behörighetsreglerna till gymnasiet 2011, varför konsekvenserna för dem inte kan skönjas i figuren.

Figur 15 Andel elever som har gått ut gymnasiet vid 19 års ålder, uppdelat på betygsdecil



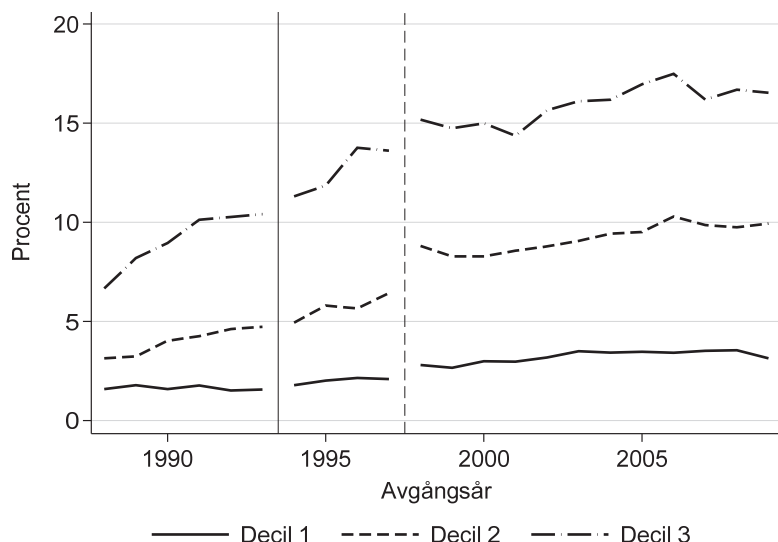
Not: Figuren visar andel avgångselever som avslutat gymnasiet vid 19 års ålder, för tiondelen av eleverna med de lägsta grundskolebetygen (decil 1), tiondelen med de näst lägsta betygen (decil 2), samt tiondelen med de tredje lägsta betygen (decil 3). Elever antas ha fullföljt sina gymnasiestudier om de antingen erhållit slutbetyg eller varit registrerade på det sista året av ett nationellt program. Den första heldragna lodräta linjen anger införandet av det mål- och kunskapsrelaterade betygssystemet på gymnasiet, den streckade linjen visar införandet av behörighetskrav till nationella program medan den andra heldragna linjen anger införandet av Gy11.

Figur 16 Andel elever som har gått ut gymnasiet vid 21-års ålder, uppdelat på betygsdecil



Not: Figuren visar andel avgångselever som avslutat gymnasiet vid 21 års ålder, för tiondelen av eleverna med de lägsta grundskolebetygen (decil 1), tiondelen med de näst lägsta betygen (decil 2), samt tiondelen med de tredje lägsta betygen (decil 3). Elever antas ha fullföljt sina gymnasiestudier om de antingen erhållit slutbetyg eller varit registrerade på det sista året av ett nationellt program. Den första heldragna lodräta linjen anger införandet av det mål- och kunskapsrelaterade betygssystemet på gymnasiet, den streckade linjen visar införandet av behörighetskrav till nationella program medan den andra heldragna linjen anger införandet av Gy11.

Figur 17 Andel som går på högskolan, uppdelat på betygsdecil



Not: Figuren visar andel avgångselever som börjat läsa på högskolan vid 25 års ålder, för tiondelen av eleverna med de lägsta grundskolebetygen (decil 1), tiondelen med de näst lägsta betygen (decil 2), samt tiondelen med de tredje lägsta betygen (decil 3). Den heldragna lodräta linjen anger införandet av det mål- och kunskapsrelaterade betygssystemet på gymnasiet, medan den streckade linjen visar införandet av behörighetskrav till nationella program.

Figur 17 visar att relativt få elever med svaga grundskolebetyg läser på högskolan.¹⁵ Det finns dock en svag uppåtgående trend under den studerade perioden. Särskilt intressant är att notera att det för elever i decil 2 och 3 finns en ökning av andelen på eftergymnasiala utbildningar i samband med de skärpta behörighetskraven till gymnasiet 1998. Det kan bero på att fler elever i dessa grupper valde att läsa ett högskoleförberedande program på gymnasiet när de målrelaterade betygen fick genomslag i gymnasievalet.

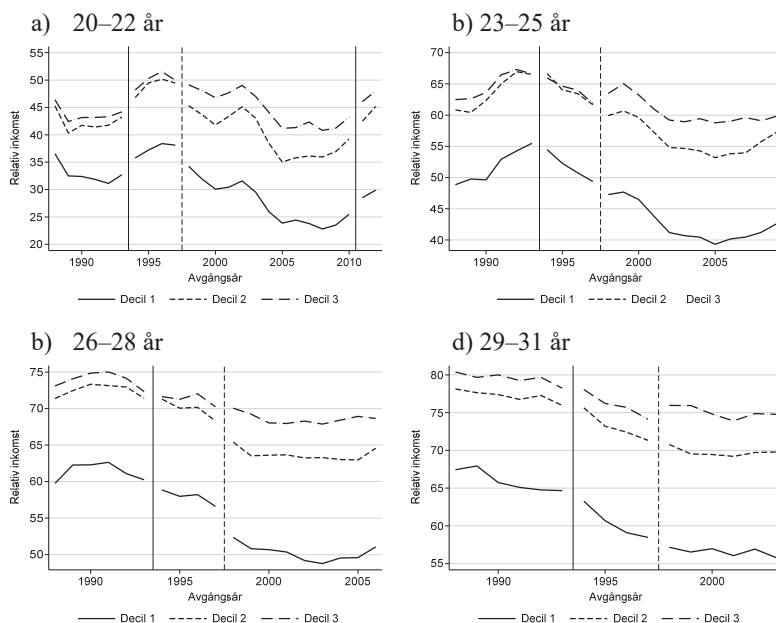
Sammanfattningsvis tycks nya betygssystem och behörighetsregler påverka gymnasieval och genomströmning, men också andelen elever som läser vidare. Det är emellertid elever med de lägsta grundskolebetygen (decil 1) som främst drabbas och begränsas av behörighetsreglerna. De hänvisas i högre grad till

¹⁵ Det sker ingen tydlig förändring av andelen elever på komvux i samband med de skärpta behörighetskraven till gymnasiet, trots att färre lågpresterande elever fullföljde sina gymnasiestudier.

individuella programmet/introduktionsprogram, vilket gör att det tar längre tid att klara av gymnasiet och att färre avslutar sina gymnasiestudier.

Vilka konsekvenser har den försämrade genomströmningen på gymnasiet för lågpresterande elevers inträde på arbetsmarknaden? I Figur 18 studeras arbetsinkomsterna vid olika åldrar. För att delvis hantera konjunktursvängningarna redovisas inkomsterna i förhållande till 45-åringars inkomster samma år.¹⁶ Delfigur a) visar att inkomsterna för 20–22-åringar tycks öka något vid övergången till målrelaterade betyg på gymnasiet, medan det framgår av delfigurer b–d att inkomstförstärkningen inte kvarstår på sikt.

Figur 18 Arbetsinkomster vid olika åldrar i förhållande till 45-åringars inkomster, uppdelat på betygsdecil



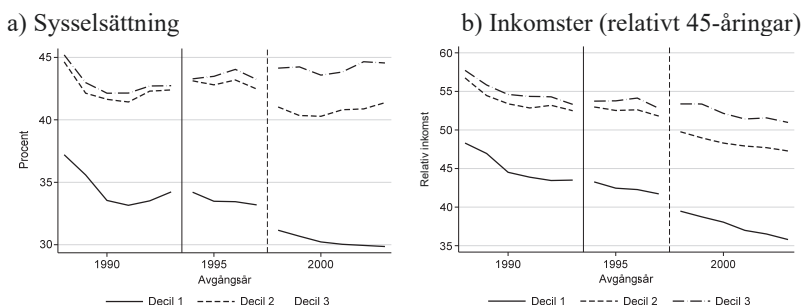
Not: Figuren visar relativ inkomst vid olika åldrar i förhållande till 45-åringar, för tiondelen av eleverna med de lägsta grundskolebetygen (decil 1), tiondelen med de näst lägsta betygen (decil 2), samt tiondelen med de tredje lägsta betygen (decil 3). Den första heldragna lodräta linjen anger införandet av det mål- och kunskapsrelaterade betygssystemet på gymnasiet, den streckade linjen visar införandet av behörighetskrav till nationella program medan den andra heldragna linjen anger införandet av Gy11.

¹⁶ Det är dock tydligt att detta inte helt hanterar konjunktursvängningarna, vilket försvårar tolkningen av utvecklingen.

För att ge en mer samlad bild av hur individernas arbetsmarknadsinträde påverkas av gymnasiereformerna, visar

Figur 19 genomsnittlig sysselsättning och inkomster vid 17–31 års ålder. Övergången till målrelaterade betyg på gymnasiet verkar bara ha haft marginell inverkan på elevernas etablering på arbetsmarknaden. Däremot tycks skärpningen av behörighetsreglerna till nationella gymnasieprogram ha försämrat försörjningsmöjligheterna för elever med de lägsta grundskolebetygen. Såväl sysselsättning som inkomster sjunker för elever i de två lägsta betygsdecilerna, medan de faktiskt förstärks något för elever i den tredje decilen. Det finns därmed inte mycket som tyder på att skärpta behörighetsregler och individuella/introduktionsprogram gynnar studiesvaga elevers gymnasiestudier eller arbetsmarknadsinträde.

Figur 19 Genomsnittlig sysselsättning och arbetsinkomster vid 17–31 års ålder, uppdelat på betygsdecil



Not: Figuren visar genomsnittlig sysselsättning och inkomst vid 17–31 års ålder, för tiondelen av eleverna med de lägsta grundskolebetygen (decil 1), tiondelen med de näst lägsta betygen (decil 2), samt tiondelen med de tredje lägsta betygen (decil 3). Inkomsterna relateras till inkomsterna för 45-åringar samma år. Den heldragna lodräta linjen anger införandet av det mål- och kunskapsrelaterade betygssystemet på gymnasiet, medan den streckade linjen visar införandet av behörighetskrav till nationella program.

5 Kausala studier och evidens från andra länder

Som visats i tidigare avsnitt uppvisar en betydande andel av eleverna låga kunskapsnivåer och riskerar därför att underkännas i ett system med centralt stipulerade gränser för godkänd nivå. Vi har också visat att övergångarna till

målrelaterade betyg på gymnasiet och i grundskolan, och införande av behörighetskrav till gymnasiet ser ut att ha påverkat de allra svagaste eleverna negativt så till vida att färre klarar gymnasiet, de blir fördröjda ut på arbetsmarknaden och har lägre inkomster. Den beskrivande analys vi presenterat kan dock inte göra anspråk på att fastställa trovärdiga orsakssamband. En central fråga är därför vad forskningen säger om de kausala effekterna av att införa godkäntränser eller behörighetsregler. Något generellt svar på denna fråga kan emellertid omöjligt ges då det beror på det system som eleverna befinner sig i. Däremot går det utifrån forskning att ge vissa svar men det är viktigt att bära med sig att dessa svar är kontextberoende, bl.a. är det ofta en kombination av regelverk som spelar in. Det är även viktigt att ha i åtanke att de elever som direkt berörs av en godkäntränser och inte klarar den är lågpresterande. Detta innebär bl.a. att de uppvisar en relativt låg etablering på arbetsmarknaden och en hög grad av sociala problem, oaktat vilket urvalssystem som används. För att isolera effekterna av skarpa behörighetsgränser krävs därför metodologiska ansatser som använder sig av en relevant jämförelsegrupp.

Det finns åtminstone tre typer av studier som är relevanta för att förstå hur en tydlig godkäntränser i betygssystemet påverkar elevers resultat på kort och lång sikt. Den första kategorin av studier försöker mäta effekterna av de incitament som ökade kunskapskrav kan medföra. De undersöker vanligen hur studieresultaten påverkas för alla elever, men inkluderar ofta även separata analyser för olika elevgrupper. En annan forskningsinriktning analyserar hur elever som på marginalen missar kunskapsmålen klarar sig i jämförelse med de som når målen, vad gäller bl.a. studieprestationer, utbildningsdeltagande, kriminalitet och inträde på arbetsmarknaden. Resultaten avser vanligen de elever som befinner sig relativt nära kravnivåerna och är därför inte lika informativa om hur elever i andra delar av fördelningen påverkas. Den sista gruppen av studier försöker mäta nettoeffekterna av att införa system där skärpta kunskapskrav kan vara en del, både vad gäller de incitament som kan uppstå och de konsekvenser som kraven kan medföra. Dessa studier är särskilt relevanta för att göra en helhetsbedömning av hur en tydlig godkäntränser i betygssystemet påverkar elevers resultat, inte minst eftersom incitamenten att förbättra resultaten vanligen är starkare ju allvarigare de negativa följdverkningarna av att missa målen är.

5.1 Studier som mäter incitamenteffekter av ökade krav

Ett vanligt argument för att införa en tydlig godkäntränser i betygssystemet är att skapa incitament för både elever och lärare att anstränga sig mer för att nå målen. Det finns också en omfattande forskning som undersöker hur betygskrav och hur

provens konsekvenser påverkar elevers inlärnin. En grupp studier undersöker hur strängheten i lärarnas bedömning påverkar elevernas prestationer. Hur hårda lärare (eller skolor) är i bedömningen mäts genom att jämföra elevers betyg med deras resultat på nationella prov, där lärare som ger lägre betyg för samma prestation anses vara strängare. (Bonesrønning 1999; 2004; Betts och Grogger 2003; Figlio och Lucas 2004; Gershenson m.fl. 2022) visar att elevers resultat förbättras om de har lärare som ställer högre krav vid bedömningen av deras prestationer. Dock finns det tecken på att framför allt högpresterande elever gynnas av hårdare betygssättning, medan avhoppen ökar i olika minoritetsgrupper.

En invändning mot dessa studier är att skillnaden mellan betyg och provresultat kan spegla andra skillnader mellan elever (såsom socio-emotionella förmågor) eller lärare (såsom undervisningsstil) som påverkar studieresultaten. En del nyare studier använder därför olika institutionella förhållanden för att fastställa orsakssamband mellan strängheten i bedömningen och elevers resultat. Bowden m.fl. (2023) studerar en reform som medförde generösare bedömning av elevers provresultat i ett amerikanskt skoldistrikt. De finner att lägre krav vid betygssättningen i genomsnitt hade små effekter på elevernas betyg (trots att de borde ha förbättrats), men att elevfrånvaron ökade. Resultaten skiljer sig dock mellan olika elevgrupper; högpresterande elever förbättrade sina betyg medan frånvaron ökade bland lågpresterande elever. Hvidman och Sievertsen (2021) studerar en reform i Danmark som ledde till att vissa gymnasieelever fick lägre betygs-genomsnitt medan andra elever inte påverkades. De finner att elever som nedgraderades – vilket kan jämföras med hårdare bedömning – presterade bättre vid senare provtillfällen. Förbättringen berodde dels på att elever som bedömdes striktare lärde sig mer, dels på att lärare kompenserade dessa elever vid senare bedömningar. En slutsats från dessa studier är att strängare krav tycks öka resultat-spridningen och att svagare grupper ofta missgynnas.

En relaterad forskningsinriktning undersöker hur elevernas prestationer påverkas av resultatens betydelse för deras fortsatta utbildningskarriär. Jacob (2005) mäter effekten av att införa tydliga kunskapskrav i grundskolorna i Chicago, både att eleverna skulle klara ett kunskapstest för att själva få avancera till nästa årskurs och för att skolorna skulle undvika strängare tillsyn. Han visar att resultaten förbättrades mer i Chicagos grundskolor jämfört med skolorna i andra liknande städer när systemet infördes. Resultatförbättringarna verkade bero på att eleverna ansträngde sig mer och att de i större utsträckning tillägnade sig och undervisades i kunskaper som var viktiga för att klara proven. Lärarna agerade också strategiskt genom att ge eleverna mer stödundervisning och fokusera på ämnen med standardiserade prov på bekostnad av andra ämnen.

Flera studier visar också att elevernas prestationer förbättras när deras resultat används vid antagningen till nästa utbildningsnivå. Bach och Fischer (2020) finner att elevresultaten försämras i tyska delstater som avskaffar provbaserad uppdelning av elever i årskurs 4, särskilt bland lågpresterande elever. Eleverna studerar mindre när kraven minskar, men deras välbefinnande verkar förbättras. Haraldsvik (2014) och Fidjeland (2023) undersöker hur införandet av betygsbaserad antagning till gymnasieskolan i flera norska län påverkar grundskoleelevers resultat. De finner att resultaten förbättras när betygen blir viktigare för antagningen till gymnasiet, särskilt för elever i områden med hård konkurrens om gymnasieplatserna där eleverna kan bredda sina valmöjligheter genom förbättrade resultat.

Rodriguez (2023) studerar gymnasieelever som precis klarar kraven för att bli behöriga till ett välrenommerat amerikanskt universitet och finner att deras studieresultat förbättras, troligen för att de försöker öka sina chanser att bli antagna. Dessa effekter är inte enhetliga över alla elevgrupper: minoritetsstudenter och mindre bemedlade studenter precis ovanför gränsen minskade sin frånvaro och drog på sig färre disciplinära åtgärder. Mer bemedlade elever ägnade sig däremot i högre grad åt strategiska kursval. Även om skarpa trösklar kan få elever att arbeta hårdare så skapas även incitament att strategiskt välja mindre krävande kurser eller mer betygsgenerösa skolor.

Att elever och familjer agerar strategiskt visas även av Cullen m.fl. (2013). De undersöker en regel som gav platsgaranti på ett selektivt delstatsuniversitet till alla som gick ut som de 10 procent främsta avseende avgångsbetygen från sin gymnasieskola.¹⁷ Cirka fem procent av eleverna som både hade motiv och möjlighet att byta skola gjorde detta som reaktion på antagningssystemet. Den ökade möjligheten att komma in på ett attraktivt universitet fick dem att välja mindre attraktiva gymnasieskolor. Hur detta strategiska beteende påverkade elevernas kunskaper är däremot oklart.

En äldre översikt av Holme m.fl. (2010) sammanfattar den då existerande forskningen kring gymnasieavgångsprov som syftar till att säkerställa att alla som får examen har en viss lägstanivå på sina kunskaper. Trots att det är dessa provs uttalade syfte finns det inte mycket som tyder på att lågpresterande elevers studieresultat förbättras av denna typ av åtgärder. Däremot finns det resultat som tyder på att förekomsten av sådana prov i sig försämrar prestationerna i en del minoritetsgrupper. I linje med detta tyder ett flertal studier på att avhoppsfrekvensen från gymnasiet ökar, särskilt bland elever på skolor i socialt utsatta miljöer.

¹⁷ Ett selektivt universitet är en institution där inte alla behöriga sökanden kommer in. I USA finns också högskolor, så kallade community colleges, där alla som är behöriga får tillträde.

5.2 Konsekvenser av att missa kunskapsmålen

Det finns en omfattande forskning som studerar effekterna av att missa uppsatta kunskapsmål under olika faser av skolgången, exempelvis att inte nå olika betygskriterier, att få underkänt vid examensprov eller att inte klara behörighetskrav. De mest trovärdiga studierna jämför elever som precis missar målen med elever som precis klarar dem.

Den sannolikt mest relevanta studien på området är Halapuu (2021). Hon undersöker hur elever som, på grund av de skärpta behörighetsreglerna i Gy11, inte blev behöriga till något nationellt gymnasieprogram. Som jämförelsegrupp används studenter med motsvarande studieresultat men som ändå klarade behörighetskraven. Resultaten visar att de skärpta behörighetsreglerna utestängde en betydande andel lågpresterande elever från de nationella programmen. Detta ledde i sin tur till förlängd gymnasieskolgång och ett kraftigt ökat inflöde till aktivitetsersättningen. Ökningen uppgick till fem procentenheter, vilket är mer än en fördubbling. Det ökade inflödet i aktivitetsersättningen berodde främst på att fler unga beviljades ersättning för förlängd skolgång. Studiens uppföljningsperiod är för kort för att kunna svara på de långsiktiga effekterna, men ungefär en tredjedel av individerna kvarstod i aktivitetsersättning efter tio år. Som Halapuu påpekar är inläsningseffekterna i aktivitetsersättningen betydande.

En annan relevant studie undersöker effekterna av tillgång till gymnasieutbildning på kriminellt beteende bland unga män i Finland (Huttunen m.fl. 2023). Författarna utnyttjar att det finns antagningsgränser i översökta program och jämför dem som precis kommer över denna gräns med dem som hamnar precis under antagningsgränsen. De finner att den som inte antas till någon gymnasieskola har en betydligt högre sannolikhet att begå brott. Inom en tioårsperiod efter antagningen är sannolikheten att män som antagits till en gymnasieskola döms för ett brott 52 procent lägre än för dem som inte antogs. Dessa effekter är störst under åren direkt efter skolantagningen och resultaten tyder på en minskad sannolikhet att någonsin begå brott, snarare än att bara fördröja brottslighetens början. Studien kan däremot inte säga något tydligt om hur arbetsmarknadsutbildningen påverkas bland dem som precis förvägrats en utbildningsplats. Det bör påpekas att elever som inte antogs alltså inte tillbringade gymnasietiden på skolbänken och alltså hade "tid" att begå brott.

I Danmark skärptes antagningsreglerna till yrkesprogram 2015 och två studier har undersökt konsekvenserna av denna reform för marginalelever, dvs. elever som hade antagits under de gamla reglerna men som inte blir behöriga under de nya. De elever som inte blev behöriga erbjöds i stället förberedande kurser och en så kallad 10:e årskurs. Föga förvånande finner Bingley m.fl. (2023)

att andelen marginalelever som gick vidare till och slutförde en yrkesutbildning minskade kraftigt. Sex år efter att dessa elever lämnat grundskolan har de väsentligt lägre sysselsättning och inkomst. Larsen m.fl. (2022) finner också en kraftig ökning av brottsligheten bland marginaleleverna efter reformen. Denna ökning är stor (två procentenheter eller en ökning på 30 procent), men till skillnad från Huttunen m.fl. (2023) verkar ökningen vara kortsiktig. Ingen av dessa studier kan studera riktigt långsiktiga effekter eftersom uppföljningsperioden är tämligen kort.

Liknande resultat står att finna för USA i en studie av Eren och Ozturk (2024) som jämför elever som hamnade precis under respektive över godkänthgränsen vid ett slutprov på gymnasienivå. De finner ökad brottslighet och ökat bidragsberoende bland minoritetselever som inte klarade proven. Detta är elever som befinner sig i en mer socialt utsatt situation än majoritetselever, vilket tyder på att den ytterligare nackdel som ett underkänt prov innebär kan drabba socialt utsatta elever särskilt hårt.

Machin m.fl. (2020) undersöker konsekvenserna av ett viktigt nationellt prov i slutet av den obligatoriska skolgången i England. De visar att studenter med likvärdiga förmågor får tillgång till avsevärt olika utbildningsvägar beroende på om de precis klarar eller misslyckas på detta prov. Tre år senare har elever som precis misslyckas med att uppnå den nödvändiga tröskeln en lägre sannolikhet att komma in på en gymnasie- eller yrkesutbildning på nästa nivå. De som inte klarar tröskeln är också mer benägna att lämna skolsystemet vid 18 års ålder, utan att ha någon form av anställning. Att inte klara provet har således viktig och potentiellt långsiktiga konsekvenser för dem som drabbas.

En norsk studie undersöker effekterna av att underkännas på ett gymnasialt slutprov på elevernas psykiska hälsa bland elever på högskoleförberedande program (Beck m.fl. 2024). För att skapa en relevant jämförelsegrupp använder de en matchningsansats som tar hänsyn till elevernas tidigare skolresultat, elevernas lärarsatta betyg och en mängd andra sociala och demografiska karakteristika. Ansatsen är inte helt övertygande då det är något icke-observerbart som får vissa elever att misslyckas på provet medan andra lyckas. Med detta sagt så finner de en tydligt förhöjd risk att få behandling för psykisk ohälsa efter att ha misslyckats.

I många skolsystem får lågpresterande elever gå om en årskurs för att på så sätt ges tid att förbättra sina resultat. Det finns en omfattande litteratur om hur elevers prestationer påverkas av att gå om en årskurs (Valbuena m.fl. 2021; Goos m.fl. 2021). De kortsiktiga effekterna beror dock på vem man jämför med. Elever som går om en årskurs förbättrar oftast sina resultat i förhållande till elever i samma årskurs (som är yngre och har gått kortare tid i skolan), men försämrar

dem i förhållande till jämgamla elever (som befinner i högre årskurser). Ett fåtal studier har kunnat studera effekterna på längre sikt och finner att elever som går om en årskurs löper större risk att hoppa av gymnasiet (Jacob och Lefgren 2009; Manacorda 2012) och har större benägenhet att begå brott (Eren m.fl. 2022).

Det finns även forskning från andra länder – och då främst USA – om effekterna av kravgränser under och efter gymnasiet. Det har redan nämnts att översikten av Holme m.fl. (2010) finner att avgångsprov som syftar till att säkerställa en viss lägstanivå vid examen från motsvarande gymnasienivå (*high school*) inte verkar leda till förbättrade resultat bland lågpresterande elever. Snarast finns det tecken på försämrade prestationer och fler avhopp, särskilt bland elever på skolor i socialt utsatta miljöer. Arbetsmarknadskonsekvenserna är oklara, vilket eventuellt kan förklaras av svårigheten att följa eleverna över tid (registerdata saknas). Vad gäller skolors agerande som svar på införande av examensprov av denna typ så finns det belägg för att skolorna ofta sätter in stödåtgärder för svaga elever, men att de har begränsade positiva effekter för eleverna. Vidare verkar tillgången till stöd att vara olikvärdig. Det finns också ett flertal studier som finner att skolor vidtar åtgärder för att skydda sig själva och sitt rykte, även med åtgärder som har negativa konsekvenser för eleverna, till exempel att på olika sätt undanta elever från att skriva prov. (Sipple m.fl. 2004; Heilig och Darling-Hammond 2008).

Papay m.fl. (2010) använder en diskontinuitetsdesign för att utvärdera effekterna av att precis misslyckas eller klara ett gymnasieavgångsprov på sannolikheten att ta gymnasieexamen i delstaten Massachusetts. För de flesta elever påverkas inte sannolikheten att ta examen av att precis misslyckas med sitt första prov i matematik eller engelska i årskurs 10. Dock har låginkomsttagare i storstadsområden som precis misslyckas med matematikprovet åtta procentenheter lägre sannolikhet att ta examen jämfört med elever som precis klarar det. Det finns ingen motsvarande effekt för att precis misslyckas eller klara provet i engelska. För dessa studenter påverkas inte sannolikheten att avancera i tid i skolan, men det leder till att de är fyra procentenheter mer benägna att hoppa av skolan året efter provet. Låginkomststudenter i urbana områden är lika benägna att göra om provet som elever i mer välbeställda bostadsområden. De har emellertid mindre framgång när de tar om provet, vilket kan bero på att de inte har samma stödresurser till sitt förfogande.

Svenska elever som inte når behörighet till gymnasiet hänvisas idag till introduktionsprogram för att bättra på sina förkunskaper. Det är svårt att kausalt uttala sig om vad effekten av detta är, men ett liknande system finns på högskolenivå på vissa håll i USA. På en del amerikanska universitet erbjuds förberedelsekurser för studenter som inte bedöms ha tillräckliga förkunskaper

för att klara av sina studier. Valentine m.fl. (2017) sammanfattar litteraturen om effekterna av att gå förberedelsekurser på högskolan och finner att studenter som gått en förberedelsekurs trots detta löper större risk att senare hoppa av sina studier än de som påbörjat studier utan att behöva gå en förberedelsekurs.

5.3 Samlade effekter av att införa kunskapskrav

Att skolor svarar på incitament kopplade till elevernas resultat har lyfts i tidigare avsnitt. En del av dessa responser kan vara positiva för eleverna men det finns gott om exempel på att skolor istället svarar strategiskt genom att exempelvis dölja skolmisslyckanden. Figlio och Loeb (2011) lyfter också fram att i den mån kunskapskraven på skolor innebär att elever faktiskt får tillgång till stöd, bättre lärare och bättre undervisning är effekter av system som håller skolor ansvariga för elevernas resultat, så kallade *accountability*-system, positiva. Däremot är det inte självklart att skolor svarar på detta sätt.

En metastudie av Redding och Nguyen (2020) undersöker särskilt kraftfulla åtgärder riktade mot lågpresterande skolor. Skolförändringar, *school turnarounds* på engelska, har vuxit fram som en strategi för att förbättra skolor med kroniskt låga resultat. Analysen sammanfattar resultaten från 35 studier som undersöker sambandet mellan dylika skolförändringar och elevernas resultat. Resultaten tyder på att åtgärderna är förknippade med ökad elevnärvaro, förbättrade resultat på standardiserade prov och högre examensfrekvens. När resultaten delas upp efter olika *turnaround*-modeller är det de som syftar till att ändra arbetssätt och skolkultur som är förknippade med förbättrade resultat.¹⁸ Däremot finner de inga belägg för signifikanta samband mellan skolstängningar eller delstatliga skolövertaganden och elevernas resultat. Detta innebär att resoluta ingripanden mot skolor med låga resultat kan bidra till förbättringar. Sådana åtgärder bör emellertid vara noga avvägda och föremål för utvärderingar. Än viktigare är att åtgärder riktade mot skolor inte på något sätt förutsätter att elever underkänns.

Det svenska systemet riktar genom godkänthetskrav och behörighetskrav de skarpa incitamenten mot elever snarare än mot skolor. Forskning som undersöker systemeffekter av ändrade krav på eleverna är därför relevanta. Med systemeffekter menar vi konsekvenser för de elever som inte direkt påverkas av kraven, exempelvis genom att undervisningen eller elevsammansättningen förändras. Denna forskning är dock betydligt mer begränsad än den som undersöker hur direkt berörda elevgrupper eller skolor påverkas. Guyon m.fl.

¹⁸ Mer specifikt är det modellerna ”*transformation*”, ”*turnaround*” och ”*restart*” som är associerade med bättre resultat. Det ligger bortom denna rapportens syfte att diskutera dessa modeller och den intresserade läsaren hänvisas till Redding och Nguyen (2020).

(2012) undersöker emellertid en relevant nordirländsk reform som kraftigt expanderade antalet platser på mer selektiva gymnasieinriktningar. Inflödet till dessa utbildningar ökade med 15 procentenheter per årskull och ökningen skedde främst bland elever med medelgoda skolresultat. Man kan alltså säga att kraven för att gå på vissa typer av selektiva program/skolor sänktes, alltså en reform i motsatt riktning mot att begränsa tillgången till nationella program. Trots den tydliga förändringen av elevsammansättningen på de selektiva inriktningarna påverkades inte resultaten på dessa. Däremot sjönk skolresultaten på de icke-selektiva inriktningarna när de förlorade många relativt starka elever. Sammantaget finns inga belägg för att expansionen försämrade kvaliteten på de selektiva inriktningarna och nettoeffekten av att sänka ingångskraven till de selektiva programmen på elevernas utbildning var tydligt positiv.

En delvis relaterad studie använder resultat från internationella undersökningar för att i en länderjämförelse undersöka effekterna på studieresultat av att tidigt dela upp eleverna på olika spår efter studieförmåga (Hanushek och Wößmann 2006). Studien är relevant för den svenska diskussionen eftersom det finns en oro för att utbildningskvaliteten kan försämrats om svagare elever släpps in på program som de kan ha svårare att klara av, men samtidigt är kontexten annorlunda då den undersöker tidig nivågruppering inom ramen för system där alla elever fortsätter sin skolgång. Resultaten visar emellertid att *alla* elevgruppers provresultat verkar påverkas negativt av tidig nivågruppering. Resultaten försämrats mest bland lågpresterande elever men verkar alltså även försämrats bland de relativt högpresterande. En liknande studie finner däremot att avskaffandet av denna typ av nivågruppering verkar ha bidragit till försämrade resultat bland barn till socioekonomiskt starka föräldrar, vilket kan bero på försämrade utbildningskvalitet för dessa elever (van de Werfhorst 2019). Det bör påpekas att resultaten i dessa studier trots sin motstridighet inte är helt oförenliga eftersom social bakgrund och studieresultat inte är samma sak. En metaanalys som undersöker en mängd studier på detta tema landar i att tidig nivågruppering verkar ha en svagt negativ inverkan på genomsnittsresultaten samtidigt som spridningen i resultaten ökar tydligt (Terrin och Triventi, 2023). Detta svarar dock inte direkt på frågan om hur senare nivågruppering, exempelvis i gymnasieskolan, genom att de svagaste eleverna, som i Sverige, nekas tillträde, påverkar utbildningskvaliteten för de elever som inte stängs ute.

En brittisk studie syftar till att undersöka hur inkomsterna och deras fördelning påverkas av ett mer eller mindre selektivt skolsystem (Burgess m.fl. 2020). Vi bedömer att studien är relevant eftersom selektionen i den brittiska studien påverkar tämligen svaga elever, vilket ju också de svenska

behörighetsreglerna gör. Resultaten är i linje med dem från litteraturen om nivågruppering. Ett mindre selektivt system är associerat med högre inkomster i nedre delen av fördelningen men lägre inkomster i den översta delen. Det ska här påpekas att det bara är bland dem med allra högst inkomster som den senare skillnaden är tydlig. Detta är i linje med att mindre selektiva system inte påverkar utbildningens kvalitet negativt för det stora flertalet av eleverna.

De studier som har presenterats varierar i relevans för svenska förhållanden med tydliga krav för godkända betyg. Med detta förbehåll i åtanke kan resultaten från studier som försöker isolera kausala effekter sammanfattas som följer:

1) Incitamenteffekter av ökade krav

Det finns stöd för att elever reagerar på de incitament som ökade krav medför genom att ägna mer tid åt studier, vilket leder till förbättrade studieresultat. Dessa effekter är vanligen större ju allvarigare de negativa konsekvenserna är av att missa målen. Samtidigt visar resultat att elever med särskilt svaga förutsättningar eller från socialt utsatta grupper och områden inte alltid svarar på incitament på samma sätt. En möjlig förklaring är att dessa elever riskerar att tappa motivationen för skolarbetet om kraven uppfattas som ouppnåeliga eller att skolorna saknar resurser för att stötta dem. Skolor svarar också på kravgränser genom att satsa mer resurser på elever som riskerar att inte uppnå kraven. Detta leder dock inte alltid till förbättrade elevresultat, och det finns tydliga tecken på att skolor, lärare och elever ibland försöker manipulera resultaten när de har incitament att göra så.

2) Konsekvenser av att missa kunskapsmål

Studier visar genomgående tydligt negativa konsekvenser av att missa uppsatta kunskapsmål som begränsar möjligheterna att studera vidare, särskilt på gymnasienivå. Risken är betydande att elever helt lämnar utbildningssystemet utan att ha någon alternativ sysselsättning. Flera studier finner också ökad kriminalitet och lägre etablering på arbetsmarknaden för elever som missar målen.

3) Effekter av att exkludera lågpresterande elever:

Det är svårt att dra tydliga slutsatser om de samlade effekterna av att stänga ute relativt stora grupper av lågpresterande elever från fortsatt utbildning. Studier av nivågrupperade skolsystem tyder dock inte på att inkludering eller exkludering av svagpresterande elever har särskilt stora effekter på utbildningens kvalitet för övriga elever.

6 Slutsatser

Allt tyder på att elevers förutsättningar för att klara skolan är normalfördelade. Detta innebär att en viss andel av eleverna inte kommer att klara en skarp gräns för godkänd nivå. Hur stor denna andel är beror på utbildningens kvalitet och vilken nivå som anses godkänd. I internationell jämförelse utmärker sig Sverige inte vad gäller andelen lågpresterande elever. Det finns också en betydande andel elever som inte uppnår tillfredsställande kunskapsresultat även i de skolsystem som enligt internationella kunskapsmätningar uppvisar mycket goda resultat. En slutsats är därför att det finns begränsade möjligheter att minska andelen lågpresterande elever genom en allmän förbättring av skolutbildningens kvalitet.

En annan väg till färre obehöriga elever är att sänka nivån för vad som anses vara en godkänd nivå. I praktiken sker detta i Sverige då flera faktorer samverkar i riktning mot en mer generös bedömning. Detta sker emellertid på ett godtyckligt sätt där de nationella provens svårighetsgrad, enskilda skolors och lärares bedömningar av provresultat, och vilken vikt dessa tillmäts vid betygssättningen spelar stor roll. En slutsats är att en opartisk bedömning av de nationella proven och en tydligare förankring av betygen mot dessa provresultat, med bibehållna krav för godkänt, skulle resultera i en markant ökning av andelen elever som underkänns i grundskolan och nekas behörighet till gymnasieutbildning.

Denna rapport visar att införandet av behörighetsgränser i den svenska skolan haft tydliga konsekvenser för främst lågpresterande elevers progression i det svenska utbildningssystemet. Vi kan visa att de elever som blir hänvisade till individuella/introduktionsprogram tenderar att bli fördröjda genom gymnasieskolan, eller inte avsluta den alls, vilket totalt sett försämrar deras position på arbetsmarknaden. För dessa elever vore det sannolikt bättre att få tillgång till yrkesutbildning under sin gymnasietid. Det vore också bättre för samhället som helhet, eftersom denna grupp behöver yrkeskunskaper för att kunna försörja sig i vuxenlivet. Att skjuta upp denna utbildningsinsats till komvux, eller tills dessa individer blir föremål för arbetsmarknadsutbildning, är ineffektivt ur ett investeringsperspektiv.

En invändning mot att ge dessa elever tillgång till yrkesprogram under gymnasietiden är att de saknar nödvändiga förkunskaper och därför skulle ha svårt att tillgodogöra sig utbildningen. En annan är att deras närvaro på utbildningarna skulle försämrat kvaliteten eller sänka statusen på yrkesprogrammen. Våra resultat visar dock att fler elever slutförde sina gymnasiestudier när de fick chansen att gå ett ordinarie program, vilket tyder på att de ändå hade tillräckliga förkunskaper (dåvarande yrkeslinjer var dock kortare och hade mindre teoretiskt innehåll). Vår tolkning av den internationella

forskningen är dessutom att det saknas stöd för tesen att några svagare elever i klassen skulle försämra resultaten för övriga elever så mycket att det motiverar att helt utestänga dem.

Svensk och internationell forskning visar att elever och skolor märkbart påverkas av de incitament som finns i systemet. Kravgränser tenderar därför i allmänhet att leda till att elever anstränger sig mer och förbättrar sina resultat. Däremot kan dylika kravgränser bidra till minskad motivation bland särskilt lågpresterande elever som upplever kraven som svåruppnåeliga och bland elever på socialt utsatta skolor. Bland dessa elever kan skarpa gränser ge upphov till uppgivenhet som försämrar redan svaga prestationer och bidrar till avhopp utan alternativt sysselsättning. Skolor svarar också på kravgränser, bl.a. genom att rikta mer resurser till elever som riskerar att inte uppnå kraven. Detta leder emellertid inte alltid till förbättrade elevresultat, och det finns tydliga tecken på att skolor, lärare och elever på olika sätt försöker manipulera resultaten när de har starka incitament att göra så. Bland skolor blandas konstruktiva försök att förbättra undervisningen med strategiska försök att dölja svaga resultat.

Forskningen visar genomgående tydligt negativa konsekvenser av att missa uppsatta kunskapsmål som begränsar möjligheterna att studera vidare, inte minst på gymnasial nivå. Risken är betydande för att eleverna helt lämnar utbildningssystemet utan att ha någon alternativt sysselsättning. Flera studier finner ökad kriminalitet, lägre etablering på arbetsmarknaden och ökat bidragsberoende. I vissa studier är dessa negativa konsekvenser stora medan de i andra är mindre. Denna rapport tyder på att de negativa konsekvenserna i Sverige är betydande.

Det är svårt att dra några tydliga slutsatser kring de samlade systemeffekterna av att relativt stora grupper av lågpresterande elever stängs ute från (eller släpps in till) fortsatt utbildning. Studier av bl.a. nivågrupperade skolsystem tyder emellertid inte på att det har några stora effekter på utbildningens kvalitet för övriga elever.

Att generalisera resultat från den internationella forskningen är svårt eftersom de incitament som olika kravgränser ger upphov till är mycket kontextberoende. Det finns också exempel på att skolutvecklingsprogram kopplade till elevernas resultat kan bidra till förbättringar av skolors undervisning och miljö. Det bör dock påtalas att sådan skolutveckling inte på något sätt kräver eller gynnas av att elever underkänns eller drabbas av andra typer av sanktioner på det sätt som det nuvarande svenska regelverket begränsar grundskoleelevers tillgång till fortsatt utbildning.

Referenser

- Bach, M., & Fischer, M. (2020). Understanding the response to high-stakes incentives in primary education. Bonn: Institute of Labor Economics. <https://www.iza.org/publications/dp/13845/understanding-the-response-to-high-stakes-incentives-in-primary-education>
- Beck, K. C., Lie, H., Røhr, B.-A., & Flatø, M. (2024). Distressing testing: A propensity score analysis of high-stakes exam failure and mental health. *Child Development*, 95(1), 242–260. <https://doi.org/10.1111/cdev.13985>
- Becker, W. E., & Rosen, S. (1992). The learning effect of assessment and evaluation in high school. *Economics of Education Review*, 11(2), 107–118. [https://doi.org/10.1016/0272-7757\(92\)90002-K](https://doi.org/10.1016/0272-7757(92)90002-K)
- Betts, J. R., & Grogger, J. (2003). The impact of grading standards on student achievement, educational attainment, and entry-level earnings. *Economics of Education Review*, 22(4), 343–352. [https://doi.org/10.1016/S0272-7757\(02\)00059-6](https://doi.org/10.1016/S0272-7757(02)00059-6)
- Bingley, P., Humlum, M. K., Kristensen, N., & Vejlin, R. (2023). Gennemførelse af en erhvervsuddannelse: To effektmålinger (Rapport VIVE). VIVE.
- Björklund, A., Fredriksson, P., Gustafsson, J.-E., & Öckert, B. (2010). Den svenska utbildningspolitikens arbetsmarknadseffekter: Vad säger forskningen? *IFAU Rapport 2010:13*. Uppsala: IFAU.
- Bonesrønning, H. (1999). The variation in teachers' grading practices: Causes and consequences. *Economics of Education Review*, 18(1), 89–106. [https://doi.org/10.1016/S0272-7757\(98\)00012-0](https://doi.org/10.1016/S0272-7757(98)00012-0)
- Bonesrønning, H. (2004). Do the teachers' grading practices affect student achievement? *Education Economics*, 12(2), 151–167. <https://doi.org/10.1080/0964529042000239168>
- Bowden, A., Brooks Rodriguez, V., & Weingarten, Z. (2023). The unintended consequences of academic leniency. <https://doi.org/10.26300/M6SA-M733>
- Burgess, S., Dickson, M., & Macmillan, L. (2020). Do selective schooling systems increase inequality? *Oxford Economic Papers*, 72(1), 1–24. <https://doi.org/10.1093/oeq/gpz028>
- Cullen, J. B., Long, M. C., & Reback, R. (2013). Jockeying for position: Strategic high school choice under Texas' top ten percent plan. *Journal of Public Economics*, 97(January), 32–48. <https://doi.org/10.1016/j.jpubeco.2012.08.012>
- Deary, I. J. (2020). *Intelligence: A very short introduction* (2nd ed.). Oxford: Oxford University Press.

- Dee, T. S., Dobbie, W., Jacob, B. A., & Rockoff, J. (2019). The causes and consequences of test score manipulation: Evidence from the New York Regents Examinations. *American Economic Journal: Applied Economics*, 11(3), 382–423. <https://doi.org/10.1257/app.20170520>
- Deming, D. J., & Figlio, D. (2016). Accountability in US education: Applying lessons from K–12 experience to higher education. *Journal of Economic Perspectives*, 30(3), 33–56. <https://doi.org/10.1257/jep.30.3.33>
- Ek, T. (2022). Resultat från nationella provet i matematik årskurs 9 läsåret 2021/2022. *Primgruppen Rapport 2022:3*.
- Eren, O., Lovenheim, M. F., & Mocan, H. N. (2022). The effect of grade retention on adult crime: Evidence from a test-based promotion policy. *Journal of Labor Economics*, 40(2), 361–395. <https://doi.org/10.1086/715836>
- Eren, O., & Ozturk, O. (2024). Student accountability, exit exams, and long-run outcomes. *Journal of Human Resources*. <https://doi.org/10.3368/jhr.0422-12277R1>
- Fidjeland, A. (2023). Using high-stakes grades to incentivize learning. *Economics of Education Review*, 94(June), 102377. <https://doi.org/10.1016/j.econedurev.2023.102377>
- Figlio, D., & Loeb, S. (2011). School accountability. In E. A. Hanushek, S. Machin, & L. Woessmann (red.), *Handbook of the economics of education* (Vol. 3, pp. 383–421). Amsterdam: Elsevier. <https://doi.org/10.1016/B978-0-444-53429-3.00008-9>
- Figlio, D. N., & Lucas, M. E. (2004). Do high grading standards affect student performance? *Journal of Public Economics*, 88(9), 1815–1834. [https://doi.org/10.1016/S0047-2727\(03\)00039-2](https://doi.org/10.1016/S0047-2727(03)00039-2)
- Flynn, J. R. (2007). *What is intelligence? Beyond the Flynn effect* (1st ed.). Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511605253>
- Fryer, R. G. (2011). Financial incentives and student achievement: Evidence from randomized trials. *Quarterly Journal of Economics*, 126(4), 1755–1798. <https://doi.org/10.1093/qje/qjr045>
- Gershenson, S., Holt, S. B., & Tyner, A. (2022). Making the grade: The effect of teacher grading standards on student outcomes. *IZA Discussion Papers*, 15556. Bonn: Institute of Labor Economics.
- Goos, M., Pipa, J., & Peixoto, F. (2021). Effectiveness of grade retention: A systematic review and meta-analysis. *Educational Research Review*, 34(November), 100401. <https://doi.org/10.1016/j.edurev.2021.100401>

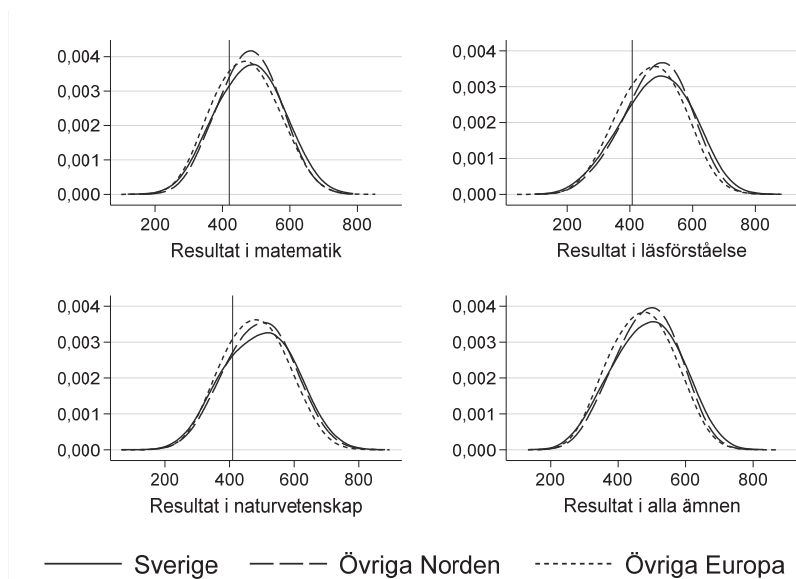
- Gust, S., Hanushek, E. A., & Woessmann, L. (2024). Global universal basic skills: Current deficits and implications for world development. *Journal of Development Economics*, 166(January), 103205. <https://doi.org/10.1016/j.jdeveco.2023.103205>
- Guyon, N., Maurin, E., & McNally, S. (2012). The effect of tracking students by ability into different schools: A natural experiment. *Journal of Human Resources*, 47(3), 684–721. <https://doi.org/10.3368/jhr.47.3.684>
- Halapuu, V. (2021). Access to education and disability insurance claims. IFAU Working Paper 2021:17. Uppsala: IFAU.
- Hanushek, E. A., & Wößmann, L. (2006). Does educational tracking affect performance and inequality? Differences-in-differences evidence across countries. *Economic Journal*, 116(510), C63–C76. <https://doi.org/10.1111/j.1468-0297.2006.01076.x>
- Haraldsvik, M. (2014). Does performance-based admission incentivize students? Oppublicerat manuskript.
- Heilig, J. V., & Darling-Hammond, L. (2008). Accountability Texas-style: The progress and learning of urban minority students in a high-stakes testing context. *Educational Evaluation and Policy Analysis*, 30(2), 75–110. <https://doi.org/10.3102/0162373708317689>
- Holme, J. J., Richards, M. P., Jimerson, J. B., & Cohen, R. W. (2010). Assessing the effects of high school exit examinations. *Review of Educational Research*, 80(4), 476–526. <https://doi.org/10.3102/0034654310383147>
- Holmlund, H., Hägglund, J., Lindahl, E., Martinson, S., Sjögren, A., Vikman, U., & Öckert, B. (2014). Decentralisering, skolval och fristående skolor: Resultat och likvärdighet i svensk skola. *IFAU Rapport 2014:25*. Uppsala: IFAU.
- Holmlund, H., Sjögren, A., & Öckert, B. (2019). *Jämlikhet i möjligheter och utfall i den svenska skolan* (SOU 2019:40). Stockholm.
- Hunt, E. (2010). *Human intelligence*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511781308>
- Husén, T. (1956). *Standardproven: En redogörelse för konstruktion och standardisering*. Stockholm: Almqvist & Wiksell.
- Huttunen, K., Pekkarinen, T., Uusitalo, R., & Virtanen, H. (2023). Lost boys? Secondary education and crime. *Journal of Public Economics*, 218(February), 104804. <https://doi.org/10.1016/j.jpubeco.2022.104804>
- Hvidman, U., & Sievertsen, H. H. (2021). High-stakes grades and student behavior. *Journal of Human Resources*, 56(3), 821–849.
- Jacob, B. A. (2005). Accountability incentives and behavior: The impact of high-stakes testing in the Chicago Public Schools. *Journal of Public Economics*, 89(5–6), 761–796. <https://doi.org/10.1016/j.jpubeco.2004.08.004>

- Jacob, B. A., & Lefgren, L. (2009). The effect of grade retention on high school completion. *American Economic Journal: Applied Economics*, 1(3), 33–58. <https://doi.org/10.1257/app.1.3.33>
- Jensen, A. R. (1998). *The g factor: The science of mental ability*. New York, NY: Praeger.
- Larsen, K. S., Andersen, L. H., & Larsen, B. Ø. (2022). Pushed out of the education system: Using a natural experiment to evaluate consequences for boys. *Journal of Experimental Criminology*, 18(2), 411–438. <https://doi.org/10.1007/s11292-020-09444-9>
- Machin, S., McNally, S., & Ruiz-Valenzuela, J. (2020). Entry through the narrow door: The costs of just failing high stakes exams. *Journal of Public Economics*, 190(October), 104224. <https://doi.org/10.1016/j.jpubeco.2020.104224>
- Manacorda, M. (2012). The cost of grade retention. *Review of Economics and Statistics*, 94(2), 596–606.
- Mansfield, J., & Slichter, D. (2024). The long-run effects of consequential school accountability. *Journal of Labor Economics*. Publicerad online. <https://doi.org/10.1086/731324>
- Neal, D. (2010). Aiming for efficiency rather than proficiency. *Journal of Economic Perspectives*, 24(3), 119–132. <https://doi.org/10.1257/jep.24.3.119>
- Neal, D., & Schanzenbach, D. W. (2010). Left behind by design: Proficiency counts and test-based accountability. *Review of Economics and Statistics*, 92(2), 263–283. <https://doi.org/10.1162/rest.2010.12318>
- Nisbett, R. E., Aronson, J., Blair, C., Dickens, W., Flynn, J., Halpern, D. F., & Turkheimer, E. (2012). Intelligence: New findings and theoretical developments. *American Psychologist*, 67(2), 130–159. <https://doi.org/10.1037/a0026699>
- Papay, J. P., Murnane, R. J., & Willett, J. B. (2010). The consequences of high school exit examinations for low-performing urban students: Evidence from Massachusetts. *Educational Evaluation and Policy Analysis*, 32(1), 5–23. <https://doi.org/10.3102/0162373709352530>
- Redding, C., & Nguyen, T. D. (2020). The relationship between school turnaround and student outcomes: A meta-analysis. *Educational Evaluation and Policy Analysis*, 42(4), 493–519. <https://doi.org/10.3102/0162373720949513>
- Riksrevisionen. (2021). *Pisaundersökningen 2018 – Arbetet med att säkerställa ett tillförlitligt elevdeltagande*. Rapport nr 2021:12. Stockholm: Riksrevisionen.

- Rodriguez, V. (2023). Student effort response to shifts in university admission policies. *Economics of Education Review*, 93(April), 102367. <https://doi.org/10.1016/j.econedurev.2023.102367>
- Ryan, R. M., & Deci, E. L. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology*, 25(1), 54–67. <https://doi.org/10.1006/ceps.1999.1020>
- Schwartz, D. L., Tsang, J. M., & Blair, K. P. (2016). *The ABCs of how we learn: 26 scientifically proven approaches, how they work, and when to use them*. W.W. Norton & Company Inc.
- Sipple, J. W., Killeen, K., & Monk, D. H. (2004). Adoption and adaptation: School district responses to state imposed learning and graduation requirements. *Educational Evaluation and Policy Analysis*, 26(2), 143–168. <https://doi.org/10.3102/01623737026002143>
- Skolöverstyrelsen. (1984). *Att välja till gymnasieskolan: Inför valet till läsåren 1984/85*. Stockholm.
- Svaleryd, H., & Vlachos, J. (2021). Skolresultat och psykisk ohälsa bland elever. Underlagsrapport till SOU 2021:89 *Sverige under Pandemin*. Stockholm.
- Terrin, É., & Triventi, M. (2023). The effect of school tracking on student achievement and inequality: A meta-analysis. *Review of Educational Research*, 93(2), 236–274. <https://doi.org/10.3102/00346543221100850>
- Valbuena, J., Mediavilla, M., Choi, Á., & Gil, M. (2021). Effects of grade retention policies: A literature review of empirical studies applying causal inference. *Journal of Economic Surveys*, 35(2), 408–451. <https://doi.org/10.1111/joes.12406>
- Valentine, J. C., Konstantopoulos, S., & Goldrick-Rab, S. (2017). What happens to students placed into developmental education? A meta-analysis of regression discontinuity studies. *Review of Educational Research*, 87(4), 806–833. <https://doi.org/10.3102/0034654317709237>
- van de Werfhorst, H. G. (2019). Early tracking and social inequality in educational attainment: Educational reforms in 21 European countries. *American Journal of Education*, 126(1), 65–99. <https://doi.org/10.1086/705500>

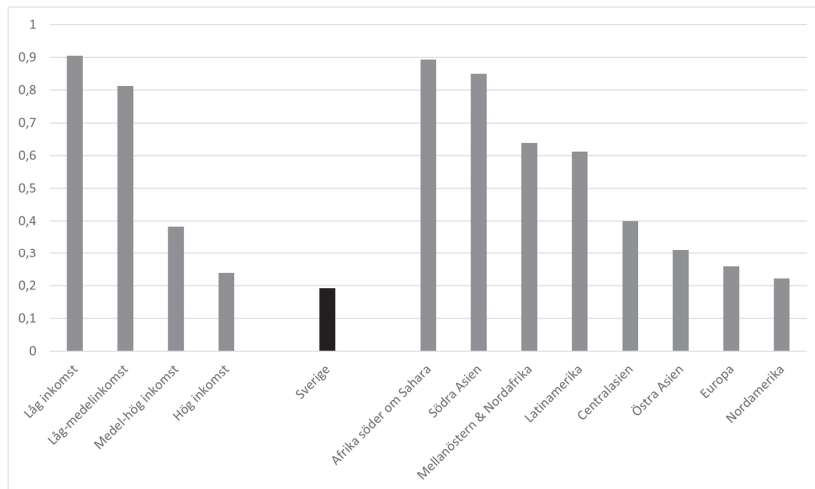
Bilaga

Figur A1 Resultat i olika ämnen för olika länder (PISA 2022)



Not: Egna analyser av data från PISA 2022. Den vertikala linjen indikerar gränsen för nivå 2 i respektive ämne.

Figur A2 Andel elever utan grundläggande färdigheter



Källa: Gust m.fl. (2024).

Prövningar i grund- och gymnasieskolan

En webbenkät om bland annat prövningar skickades under hösten 2024 till Sveriges Lärares undersökningspanel. Urvalet omfattade 2 000 lärare i grundskolan årskurs 7-9 och lärare i gymnasieskolan. Insamlingen pågick mellan 27 september till 11 oktober 2024 och två påminnelser skickades. Svarsfrekvensen blev 42 procent. En utförligare metodbeskrivning återfinns i slutet av denna PM.

Undersökningsresultat prövningar

Diagram 1. Har du någon gång under de tre senaste åren varit involverad i prövningar med elever?

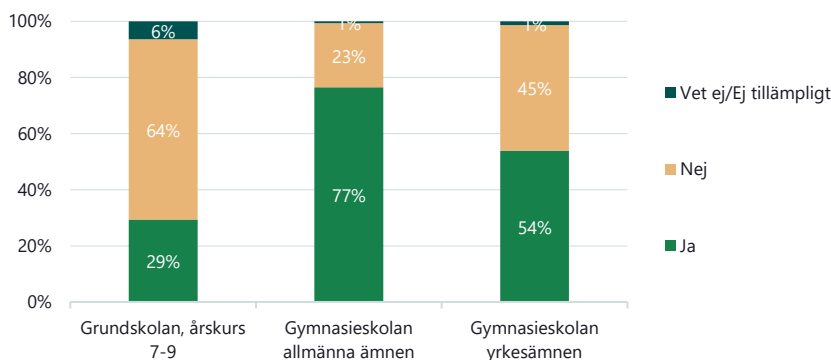
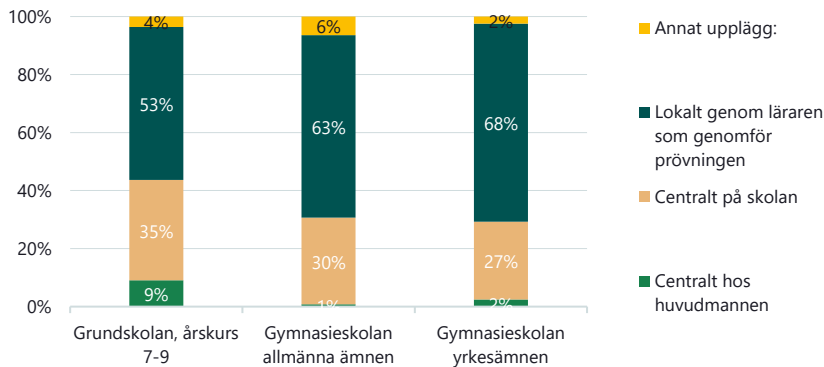
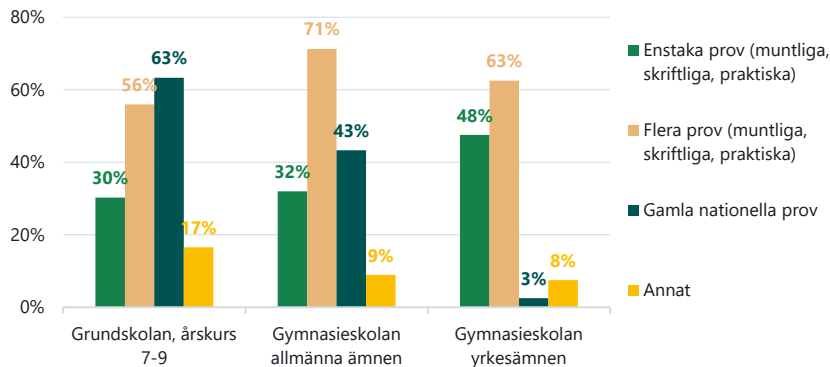


Diagram 2. Vem organiserar arbetet med provningar?



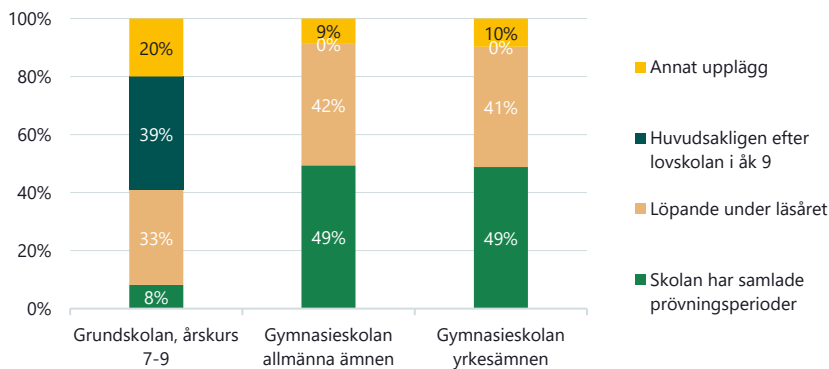
Annat: "Ämneslaget, arbetslaget, gemensamt mellan flera"

Diagram 3. Vad innehåller provningar i dina ämnen?



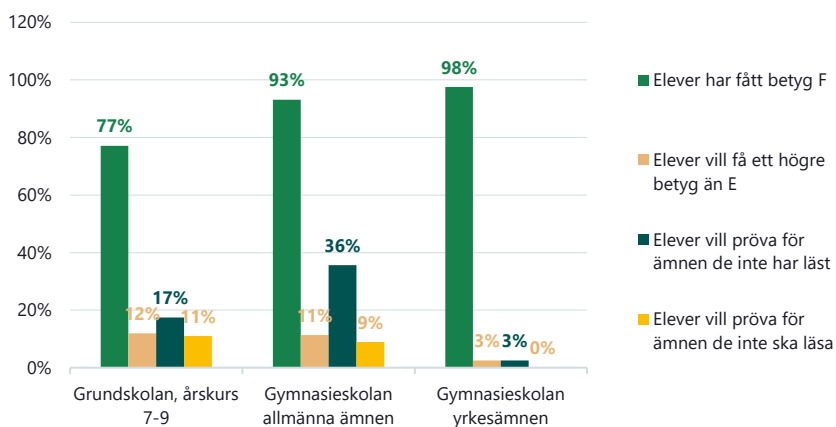
Annat: "gamla kursprov, arbetsprov, inlämnade texter, redovisningar etcetera. Svårt att genomföra provningar i estetiska ämnen."

Diagram 4. Hur ofta sker prövningar vanligtvis?

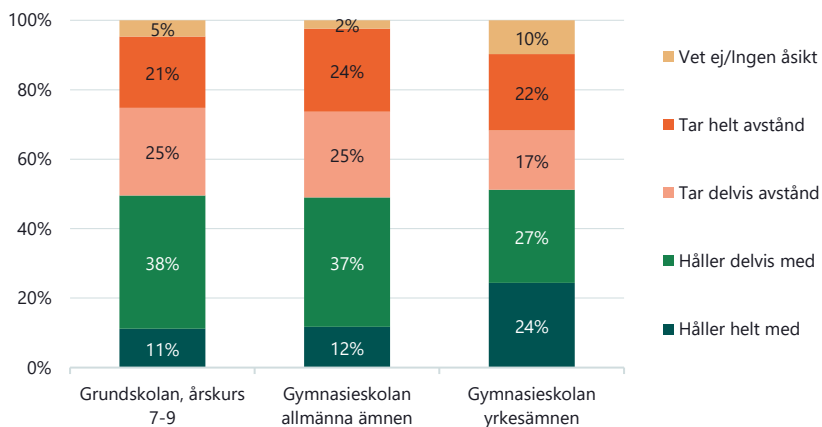


Annat: "Vid behov, det varierar, sällan"

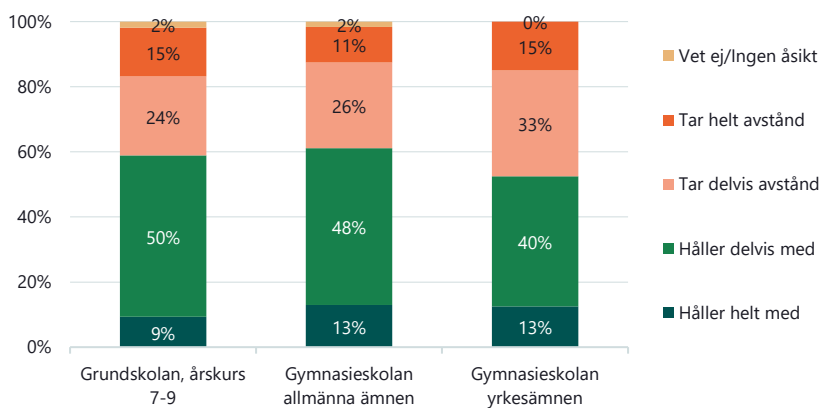
Diagram 5. Av vilket/vilka skäl har prövningarna i dina ämnen skett under de senaste tre åren?



**Diagram 6 I vilken utsträckning håller du med om följande påstående:
Organisationen kring prövningar fungerar bra**



**Diagram 7: I vilken utsträckning håller du med om följande påstående:
Prövningarna ger en korrekt bedömning av elevernas kunskaper**



**Diagram 8: I vilken utsträckning håller du med om följande påstående:
Prövningar görs alltid på hela ämnet/ämnena**

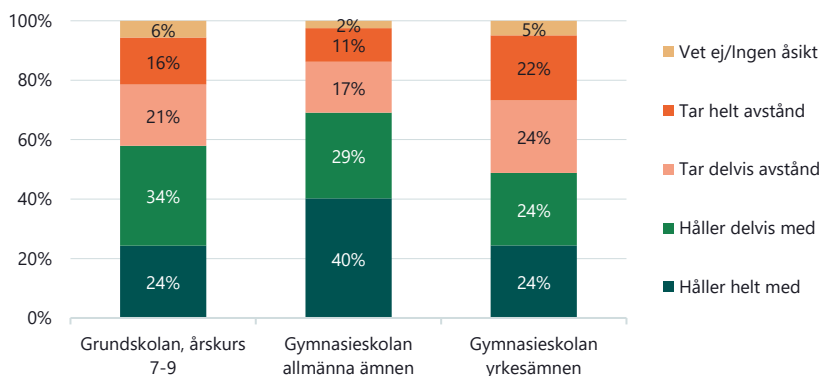
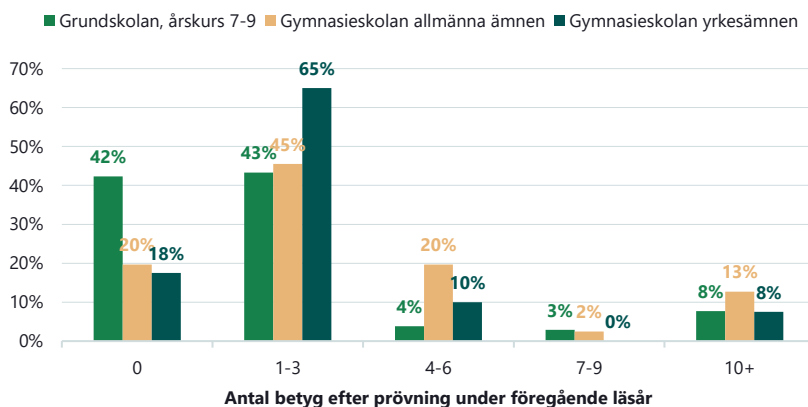


Diagram 9: Ungefär hur många betyg har du satt efter prövning under läsåret 23/24?



I genomsnitt har lärarna satt 3,2 betyg efter prövning under föregående läsår. Det högst förekommande antalet är 40 betyg.

Metodbeskrivning

En webbenkät skickades till ett representativt urval i Sveriges Lärares undersökningspanel, omfattande 2 000 lärare i skolväsendet. Av dessa arbetade 1 000 lärare i grundskolans årkurs 7-9 och resterande i gymnasieskolan. Svarsfrekvensen blev 42 procent, vilket framgår av tabell M1. Fördelningen av de svarande i frågorna gällande prövningar blev som mest 374 lärare i grundskolan, 324 lärare i gymnasieskolans allmänna ämnen och 76 lärare i gymnasieskolans yrkesämnen.

Tabell M1: urval och svarsfrekvens	Antal	Andel
Bruttourval	2 000	100%
Felaktiga kontaktuppgifter	26	1%
Ej i målgrupp	2	0%
Nettourval	1 972	100% (99%)
Antal svarande	831	
Antal fullföljda	777	
Svarsfrekvens beräknat på nettourval		42%

Insamlingsperioden var mellan 27 september och 11 oktober 2024. Totalt skickades två påminnelser. Undersökningen genomfördes inom ramen för Sveriges Lärares undersökningspanel. Till skillnad från andra paneler är Sveriges Lärares medlemspanel inte självrekryterad. Alla yrkesverksamma medlemmar har därför samma sannolikhet att ingå i panelen. Panelen upprättas en gång varje år. Sveriges Lärare organiserar en mycket stor majoritet av skolväsendets legitimerade och behöriga lärare och undersökningspanelen har mycket god representativitet, vilket framgår av tabell M2. Genom samkörningar av Sveriges Lärares medlemsregister, andra registeruppgifter och Skolverkets officiella statistik genomför Sveriges Lärare noggranna bortfallsanalyser för att säkerställa resultatens validitet och reliabilitet.

Tabell M2: Undersökningspanelens representativitet per skolform

Skolform	Tjänstgörande (Skolverket)	Sveriges Lärares undersökningspanel
Fritidshem	4%	5%
Förskola	24%	27%
Förskoleklass	4%	3%
Grundskola inkl. anpassad	49%	45%
Gymnasieskola inkl. anpassad	16%	15%
Komvux inkl. sfi och anpassad	4%	4%

Anm: Tjänstgörande personal 2022/23 med pedagogisk högskoleexamen enligt Skolverkets statistik

Statens offentliga utredningar 2025

Kronologisk förteckning

1. Skärpta krav för svenskt medborgarskap. Ju.
2. Några frågor om grundläggande fri- och rättigheter. Ju.
3. Skatteincitament för forskning och utveckling. En översyn av FoU-avdraget och expertskatte-reglerna. Fi.
4. Moderna och enklare skatteregler för arbetslivet. Fi.
5. Avgift för områdessamverkan – och andra åtgärder för trygghet i byggd miljö. LI.
6. Plikten kallar! En modern personalförsörjning av det civila försvaret. Fö.
7. Ny kärnkraft i Sverige – effektivare tillståndsprövning och ändamålsenliga avgifter. KN.
8. Bättre förutsättningar för trygghet och studiero i skolan. U.
9. På språklig grund. U.
10. En förändrad abortlag – för en god, säker och tillgänglig abortvård. S.
11. Straffbarhetsåldern. Ju.
12. AI-kommissionens Färdplan för Sverige. Fi.
13. En effektivare organisering av mindre myndigheter – analys och förslag. Fi.
14. En skärpt miljöstraffrätt och ett effektivt sanktionssystem. KN.
15. Stärkta drivkrafter och möjligheter för biståndsmottagare. Volym 1 och 2. S.
16. Ett nytt regelverk för uppsikt och förvar. Ju.
17. Anpassning av svensk rätt till EU:s avskogningsförordning. LI.
18. Ett likvärdigt betygssystem. Volym 1 och 2. U.

Statens offentliga utredningar 2025

Systematisk förteckning

Finansdepartementet

Skatteincitament för forskning och utveckling. En översyn av FoU-avdraget och expertskatte- reglerna. [3]

Moderna och enklare skatteregler för arbetslivet. [4]

AI-kommissionens Färdplan för Sverige. [12]

En effektivare organisering av mindre myndigheter – analys och förslag. [13]

Försvarsdepartementet

Plikten kallar! En modern personal- försörjning av det civila försvaret. [6]

Justitiedepartementet

Skärpta krav för svenskt medborgarskap. [1]

Några frågor om grundläggande fri- och rättigheter. [2]

Straffbarhetsåldern. [11]

Ett nytt regelverk för uppsikt och förvar. [16]

Klimat- och näringslivsdepartementet

Ny kärnkraft i Sverige – effektivare tillståndsprövning och ändamålsenliga avgifter. [7]

En skärpt miljöstraffrätt och ett effektivt sanktionssystem. [14]

Landsbygds- och infrastrukturdepartementet

Avgift för områdessamverkan – och andra åtgärder för trygghet i byggd miljö. [5]

Anpassning av svensk rätt till EU:s avskogningsförordning. [17]

Socialdepartementet

En förändrad abortlag – för en god, säker och tillgänglig abortvård. [10]

Stärkta drivkrafter och möjligheter för biståndsmottagare Volym 1 och 2. [15]

Utbildningsdepartementet

Bättre förutsättningar för trygghet och studiero i skolan. [8]

På språklig grund. [9]

Ett likvärdigt betygssystem Volym 1 och 2. [18]